

表10 バスク地区選択医療技術  
(スペイン, 1996)

---

終末期保健医療  
第一線医療での抗生物質  
慢性疾患の定期的管理  
エイズ患者の入院・外来計画  
日帰り手術  
更年期管理  
救急医療における検査・画像診断  
在宅医療  
第一線医療のルーチン検査  
麻酔前情報  
MRI  
保健医療提供者の訓練

---

非－無作為化比較試験の活用可能性に関する事例検討  
－その 1：RCT と非 RCT の結果の差異－

主任研究者 久繁哲徳 徳島大学医学部衛生学教授  
片山貴文 徳島大学医学部衛生学助手  
三笠洋明 徳島大学医学部衛生学講師

研究要旨 保健医療の評価・選択の基準には、無作為化比較試験（RCT）が用いられている。RCT は妥当性が高く、有効性を評価する上で強力な研究設計である。しかしながら、主要な問題点として、対象および保健医療が特殊な条件下で実施されるため、評価結果の普遍性が限定されることが指摘されている。そこで、非 RCT の利用可能性を、RCT と比較する中で、検討を行なった。その結果、非 RCT の結果が、RCT に比べて有意に大きい場合は、半数弱認められた。また、非 RCT の結果が RCT に比べて大きい傾向が、半数強に認められた。ただし、両者の差には、関連する要因に一定の傾向は認められなかった。とくに、非 RCT と RCT に差が認められなかった場合は関連要因の調整が十分に行なわれており、差が大きい場合は調整が行なわれていない傾向が認められた。今回の結果はパイロット的な検討であったが、その内容から考えて、今後さらに、問題をより詳細に検討することに、重要な意味があると考えられる。

目的

E B（evidence-based）接近法が、保健医療の評価・選択の基準となっている<sup>1,2)</sup>。ここでは、根拠の質の判断基準として、無作為化比較試験（RCT）が基準として用いられている。RCT の研究設計は、介入の有無を無作為に割り付けて比較するところに特徴があり、それにより比較する 2 つの群の間で、介入以外の特徴に差が無いことを保証する。

しかしながら、RCT の主要な問題点として、対象および保健医療が特殊な条件下で実施されるため、評価結果の普遍性が限定されることが指摘されている<sup>3-5)</sup>。その意味では、複数の RCT を統合したメタ・アナリシスで

も、同様な問題点が存在しており、最近では、できるだけ一般的な対象を含むように、メガ・トライアルによる評価の実施が望まれている<sup>6)</sup>。ただし、メガ・トライアルについても、実際の保健医療の状況とは、対象および介入の種類など、さまざまな問題が残されていることは明らかである。その意味では、実際の保健医療の状況を反映させた結果研究（outcome research）について、もう一度見直しをすることが必要と考えられる<sup>7,8)</sup>。

とくに、わが国の場合、わが国で実施された RCT の情報は極めて限定されている。その意味では、実施可能性の高い非 RCT について、その情報が RCT の代わりに、どこま

で利用可能なのかを検討することは重要な意味がある。そこで、今回、RCTと非RCTの研究結果を比較する中で、非RCT活用可能性の検討を行ないたいと考えた。

## 方法

### 1) RCTと非RCTの妥当性に

保健医療に関するRCTと非RCTの研究結果について、その妥当性を比較検討する際、表1に示すような問題が指摘されている<sup>3,4)</sup>。

まず内部妥当性 (internal validity) について見ると、選択の偏り (selection bias) は、対象者の選択の際に生じる。RCTでは、対象者を無作為割り付けを行っているため、選択の偏りは生じにくい。また、攪乱 (confounding) の影響も減少させることができる。一方、非RCT (ここでは、コホート研究のみを想定) では、これらの偏り、攪乱の影響が無視できない。そのために、統計学的調整 (adjustment) や対応 (matching) を行っているが、それだけでは必ずしも十分な対策とはいえない。

また、情報の偏り (information bias) は、結果の評価の際に生じる。RCTでは、薬剤の場合には、目隠し (blinding) を実施することが多いため、偏りの生じる可能性は少ないが、それ以外の介入では必ずしも十分に行われているとは言えない。また、コホート研究でも同様な状態にあり、RCTと非RCTとの間で、明確に差異があるとは言えない。外部妥当性について見ると、対象者の選択については、包含基準と除外基準が関係する。いずれもRCTでは、さまざまな条件を設定するため、極めて限定された特殊な患者が選ばれる可能性が高い。一方、非RCTでは、実際の保健医療の場に近いため、幅広い母集団を代表する可能性が高い。ただし、どのような状況から選択されているか、その特定が重要な問題となる。また、実際の参加については、RCTでは、患者の選好が重要な役割

を果たす。ある報告では、その段階で対象者が約半数近くまで減少することが指摘されている。さらに、RCTでは保健医療の内容そのものが、一般の状態とは異なる特殊な制限 (内容、種類、数ともに) が加えられるため、この点でも多くの問題が残される。

今回の検討では、まず内部妥当性に焦点を当てて検討を行った。従来、非RCTがRCTに比べて結果を過大評価する危険性が高いことが指摘されている<sup>5)</sup>。ただし、これらは、必ずしも総合的な評価に基づくものではない。その意味では、現在の状態がどのようなものであるかを検討することが求められる。

とくに、わが国ではRCTの実施状況は極めて限定されている。したがって、根拠の質の高い情報を利用しようとする、諸外国のRCTを利用することが前提となり、それを非RCTで補うことが必要となる。そこで、RCTと非RCTの最終的な結果について、その効果の規模と方向性について比較検討するとともに、内部妥当性に影響する要因の観点から問題点の吟味を行いたいと考えた。

### 2) 事例検討の方法

RCTと非RCTの結果を比較するために、MEDLINEを用いて文献検索を実施した。検索には、randomized controlled trial, research design, observational, cohort, non random, quasi experiment, nonexperiment, などの用語を組合わせて用いた。包含基準としては、1) RCT (あるいは複数のRCTs) の結果と非RCTの結果とを比較していること、2) 両者の介入は同一であること、3) 対照群の治療も同一であること、4) 結果の指標が同一であることを用いた。検索期間は1995年から1998年の4年間とした。35論文が把握できたが、上記の条件を満たすのは5件<sup>10-14)</sup>のみであった。

把握されたこれらの論文について、RCTと非RCTについて、介入の効果の大きさ、

方向性を比較検討した。さらに、非RCTにおいて、内部妥当性に関連する選択の偏りを補償するための対策についても併せて検討を行った。

## 結果

比較検討した5論文について、RCTと非RCTの研究結果を比較すると、表2に示すように、両者の間で統計学的な有意差が認められたのは、2論文であった(40%)。これらは、いずれも非RCTの効果が大きかった。

RCTと非RCTの比較結果を表3に示した。まず、両者の間で有意な差が認められた2研究について詳細を以下に示した。

Pyoralaら(1995)<sup>10)</sup>は、cryptorchidismに対するLHRHおよびhCGのホルモン療法の有効性について吟味を行った。11のRCTと22の非RCTを把握し、それぞれを併合して比較検討を実施した。結果の指標には、成功率(検査結果を利用)を用いている。その結果、RCTと非RCTとの間には、有意な差が認められた。選択の偏りを調整するための統計学的な分析については、情報は提示されていなかった。

Emanuel(1996)<sup>11)</sup>は、終末期医療について、ホスピスと通常医療との比較検討を行っている。ただし、結果の指標は、臨床的なものではなく、経済的(費用)なものである。その結果、非RCTの費用削減が、RCTに比べて有意に多かった。統計学的調整に関しては、年齢、性、癌の種類、医療サービスが実施されている。ただし、こうした調整による変動への寄与は5%程度と指摘されている。

一方、RCTと非RCTとの間で差が認められなかったのは、つぎの3研究である。

Jhaら(1995)<sup>12)</sup>は、心血管疾患の予防について、抗酸化ビタミンの効果について吟味を行っている。RCTと大規模なコホート研究をそれぞれ一つづつの比較検討を行った。指標としては心血管死亡率の相対危険の減少

を用いている。非RCTの減少が大きかったが、有意差は認められなかった。統計学的調整については、年齢、喫煙、飲酒、閉経、ホルモン利用など多数の項目について実施している。

McKayら(1995)<sup>13)</sup>は、アルコール依存症のリハビリテーションの効果について検討を行った。日中外来と入院リハビリの2種類の介入について、飲酒日数を指標として比較を行っている。その結果、RCTの日数が少なかったが、有意差は認められなかった。また、統計学的な調整は実施されていない。

Schmoorら(1996)<sup>14)</sup>は、乳癌の化学療法、ホルモン療法について検討を行っている。指標としては、無病生存率を用いている。その結果、RCTと非RCTとの間には、統計学的な有意差は認められず、ほとんど同様な結果が認められた。統計学的調整には、Coxのモデルが利用されていた。

以上の結果が示すように、RCTと非RCTとの間に有意差が認められたのは40%であったが、非RCTの結果がRCTに比べて大きい傾向は60%に認められた。ただし、両者の差が、どのような要因に関連しているか、一定の傾向は認められなかった。なお、非RCTとRCTに差が認められなかった場合は関連要因の調整が十分に行なわれおり、差が大きい場合は調整が行なわれていない傾向が認められた。

このように、RCTと非RCTについて、従来指摘されてきたことが該当する場合もあるが、必ずしも典型的に認められる例は少ないことが分かる。とくに、統計学的な調整を多変量解析で実施した場合には、両者の差が低減する可能性が推定される。

なお、今回はパイロット的な検討であったため、さらに詳細な検討が必要と考えられる。とくに、比較論文の範囲と量をさらに拡大し、さらに元の研究の詳細な比較検討が望まれる。その際、内部妥当性の問題だけでなく、外部

妥当性についても併せて検討を行い、RCTと非RCTの利害得失について評価することが求められよう。

#### まとめ

RCTと非RCTの結果を比較した結果、つぎのような結論が認められた。

- 1) RCTと非RCTの研究結果の比較検討は、極めて限定されていた。
- 2) 非RCTの結果が、RCTに比べて有意に大きい場合は、半数弱であった。
- 3) 非RCTの結果がRCTに比べて大きい傾向が、半数強で認められた。
- 4) 両者の差には、関連する要因に一定の傾向は認められなかった。
- 5) 非RCTとRCTに差が認められなかった場合は関連要因の調整が十分に行なわれおり、差が大きい場合は調整が行なわれていない傾向が認められた。

今回の結果はパイロット的な検討であったが、その内容から考えて、今後さらに、問題をより詳細に検討することに重要な意味があると考えられた。

#### 文献

- 1) 久繁哲徳，監訳：サケットら，根拠に基づく医療，薬業時報社，1998
- 2) 久繁哲徳，監訳：ミュア－グレイ，根拠に基づく保健医療，薬業時報社，1999
- 3) 久繁哲徳：臨床情報のチェックポイント，医歯薬出版，1994
- 4) Feinstein AR: Clinical Epidemiology, WB Saunders, 1985

- 5) Davis CE: Generalizing from clinical trials, Controlled Clin 15:11-14, 1994.
- 6) Borzak S, Ridker PM: Discordance between meta-analyses and large-scale randomized controlled trial, Ann Intern Med, 123:873-877, 1995
- 7) Roper WL, et al: Effectiveness in health care, NEJM, 319:1197-1202, 1988
- 8) Black NA: Why we need observational studies to evaluate the effectiveness of health care, BMJ, 312:1215-1218, 1996
- 9) Chalmers TC, et al: Bias in treatment assignment in controlled clinical trials, NEJM, 309:1358-1361, 1983
- 10) Pyorala S, et al: A review and meta-analysis of hormonal treatment of cryptorchidism, J Clin Endocrinol Metab, 80:2795-2799, 1995
- 11) Emanuel EJ: Cost-savings at the end of life. What do the data show? JAMA, 275:1907-1914, 1996
- 12) Jha P, et al: The antioxidants vitamins and cardiovascular disease, Ann Intern Med, 123:860-872, 1995
- 13) McKay JR, et al: Effect of random versus nonrandom assignment in a comparison of inpatient and day hospital rehabilitation for male alcoholics, J Consult Clin Psychol, 63:70-78, 1995
- 14) Schmoor C, et al: Randomized and non-randomized patients in clinical trials, Stats Med, 15:236-271, 1996

表1 RCTと非RCTとの比較

評価指標	RCT	非RCT
内部妥当性		
選択の偏り	小	大
情報の偏り	?	?
外部妥当性		
包含基準	狭	広
除外基準	狭	広
実際の参加	少	多

表2 RCTと非RCTとの差

差が認められる	2
差は認められない	3
総計	5

表3 RCTと非RCTの研究結果の比較

研究	疾患	介入	結果	調整	RCT	非RCT	
差あり							
Pyorala et al (1995)	Cryptorchidism	ホルモン療法	成功率	不明	LHRH	21% (18-24)	47% (43-50)
					hCG	19% (13-25)	33% (31-35)
Emanuel (1996)	終末期	組織化	費用削減	年齢, 他	3%	34%	
差なし							
Jha et al (1995)	心血管疾患	予防 (臓器)	死亡 (CVD)	年齢, 他	RR	2% (-8-11)	31% (3-51)
McKay et al (1995)	アルコール依存症	組織化	飲酒日	なし	外来	2.8±4.6	4.5±7.6
					入院	6.7±7.2	7.0±9.0
Schmoor et al (1996)	乳癌	薬剤療法	無病生存率	予後因子	6CMF	.9 (.7-1.2)	.9 (.6-1.4)
					Tamoxifen	.8 (.5-1.0)	.5 (.3-.8)
					6CMF+RT	.8 (.5-1.3)	.8 (.4-1.5)

臨床診療ガイドライン (clinical practice guideline) の  
開発方法に関する研究

主任研究者 久繁哲徳 徳島大学医学部衛生学講座教授

研究要旨 わが国においては、さまざまな形でさまざまな機関から診療ガイドライン (PG) が公表されているが、その科学性には大きな問題が残されている。そこで、科学的なPGの開発方法について検討を行った。文献の批判的吟味により、開発方法としては、非公式的合意形成、公式的合意形成、根拠に基づくガイドライン開発の3種類があること、しかも根拠に基づくガイドラインが望ましいことを明らかにした。こうした開発方法の、具体的な内容と過程とを、事例に基づき提示し、その意味について検討を行った。また、PG開発の費用についても情報を提示した。これらの検討結果から、今後のPG開発の課題を明らかにした。

目的

診療ガイドライン (practice guideline, PG) は、予防から診断、治療、リハビリまで、診療内容と様式について、専門的に要約した勧告である<sup>1, 2)</sup>。通常、「特定の臨床状態に対する適切な保健医療について、臨床医と患者の判断を支援するために、系統的に開発された声明」と定義されている<sup>1, 2)</sup>。

ガイドラインは、古くはギリシャ時代から存在するが、現在では、より厳密で科学的に検証されたものへと大きく変化している<sup>1-4)</sup>。とくに、この5年間は、現代的なPGが数多く開発され、臨床の場で広く利用されるようになってきた。その背景として、表1に示すような、医療のバラツキ、不適切な医療、医療の不確実の3つの問題が指摘されている<sup>1, 2)</sup>。しかも、医療費の高騰と一般経済の悪化は、これらの改善だけでなく、経済的効率もPGの基準として組み入れることを不可欠にしている。こうした点から、PGの主な目的

が、表2にのように要約して示されている<sup>1, 2)</sup>。PGは、基本的に患者と臨床医の臨床判断を支援することにあるが、さらにその他にも付随する主な目的がいくつか挙げられる。

近年の保健医療改革の動向から、PGはさまざまな領域において、極めて重要な位置を占めている<sup>5)</sup>。例えば、保健医療改革の中で注目を浴びている「根拠に基づく医療」は、医療テクノロジー・アセスメントにより作成した根拠を臨床行為に適用するものであるが、PGは最新で最善の根拠を要約した実践的な枠組みを提供する。また、医療組織として効果的で効率的な医療を提供するための「医療の質の改善」では、とくにケア・マネジメントの管理方法としてPGは不可欠であり、クリティカル・パスなどの中心を占める。さらに、地域の全人口集団を対象とし特定の疾患に注目する「疾病経営管理」では、PGは、医療機関の連携を進め、継続的で統合的な医療を実現するための手段を提供する。

このように、PGは現在の保健医療において重要な意味を持っているが、その質が低い場合には目的の達成はほとんど期待できないと考えられる。したがって、PGの開発の段階から適切な方法を用いることが重要な意味を持つ。PGの開発方法には、非公式の合意形成から、公式の合意形成、根拠に基づく方法、明示的な方法までさまざまなものがある。その中でも、根拠に基づく方法が最も望まれている。

わが国においては、さまざまな形でさまざまな機関からPGが公表されているが、そのほとんどは非公式の合意形成によるものであり、その科学性には大きな問題が残されている。したがって、わが国におけるPGの開発については、根拠に基づくPGの開発が重要な課題となると考えられる。

そこで、本研究では、根拠に基づくPGの開発方法について検討するとともに、その問題点と今後の課題を明らかにしたいと考えた。そして、それらの成果に基づき、わが国でのPG開発への根拠に基づく方法の適用について検討を行った。

## 方法

PGの開発(development)方法に関する情報を把握するために、MEDLINEによる文献検索を実施した。検索戦略としては、1.「clinical guideline」or「practice guideline」、2.「development」、3.「evidence」を用い、1 and 2 and 3とした。検索期間は、1996年から1998年10月とした。把握された文献は49件であった。また、1984年から1995年については、3の項目を除いて検索を行った。というのも、当時は、evidence-based medicineの概念が確立されていないため、適切な情報入手が困難なためである。その結果、33件の文献が把握できた。これら82論文とともに、手作業検索を実施し、それ以外の書籍、grey report、論文から関連情報の把握を行

った。

上記により把握した情報に基づき、PG開発方法の種類、内容、費用、問題点など主要な項目について検討を行い、わが国におけるPG開発のあり方を明らかにした。

## 結果と考察

### 1) ガイドライン開発方法の種類と特徴

#### 1. ガイドライン開発の国際的な動向

現在までに提案されているPGの数は、すでに数万を越えており、その種類と内容を把握するのは、ほとんど困難な状況である。

米国は、国際的にPGに関する活動が最も盛んな国である。中でも、予防サービス特別委員会(USPSTF)は、1989年に予防領域で包括的な記念碑PGを勧告している<sup>8)</sup>。また、臨床領域では、保健医療政策研究局(AHCPR)が本格的なPGを18作成しており<sup>9)</sup>、その他、内科学会、小児科学会を始めとする各種専門学会も、PGの開発と利用に取り組んでいる<sup>9)</sup>。また、医師会も、診断検査技術評価プログラム(DATTA)の中で勧告を出しており、PGについても内部あるいは外部との協力を積極的に取り組んでいる<sup>9)</sup>。

一方、カナダは、上記の米国のUSPSTFの基礎となる、カナダ定期健康診断特別委員会(CTSPHE)が1979年に、世界に先駆けて根拠に基づく予防のPGを報告している<sup>3)</sup>。さらに、その後も、米国とともにPGに関する活動は盛んであり、政府、医師会、学会などが協同して、1400種類のPGが収集され、データ・ベース化されている<sup>8)</sup>。また、その内200種類についてはインターネットで全テキストが入手可能である<sup>8)</sup>。こうしたことを背景に、現在は、PGの開発から、その質を評価し実行するための課題へと転換してきている。

英国では、1990年代の保健医療改革の中で、テクノロジー・アセスメントとEBMが急速



に進められており、その中で、PGの開発と利用が進められている。現在、英国でもすでに約2000種類近くのPGが存在している。そのため、米国、カナダと同様に、PGの評価と利用が注目されている。そうした課題に対応するために、PGのPGが検討されている<sup>9)</sup>。

その他、ヨーロッパの諸国でもPGの活動は同時代的な大きな波となっている。例えば、フランスでもANDEM (National Agency for the Development of Mediccal Evaluation) が中心となってPGの開発を行っている<sup>10)</sup>。

## 2. ガイドライン開発方法の種類と内容

PGの開発するための方法については、開発する組織、その組織が持つ哲学的な背景によって、大きく異なっている。しかしながら、開発方法について、系統的に検討した研究は限られている。その中から、代表的な分類として、Woolf の分類<sup>11)</sup>を用い、過去のPG開発方法の種類を提示するとともに、その内容を概括した。主なPG開発方法は、表3に示すように、1) 非公式的合意形成 (informal consensus development), 2) 公式的合意形成 (formal consensus development), 3) 根拠に基づく指針開発 (evidence-based guideline development) の3種類である。以下に、その内容を要約して示す。

### (1) 非公式的合意形成

PGは、従来より「専門家の意見」(expert opinion) に基づくものがほとんどであったし、現在も、その状況は変わっていない<sup>11, 12)</sup>。この方法の特徴は、専門家が意見を交換し、主観的に重要と考えたものを判断の根拠にすることにある。したがって、結果の指標選択、根拠の把握や評価、評価の統合には、ほとんど明確な基準と記述がともなわない。PGの内容も、ほとんどが勧告を記述するに

留まる。

表4に示すように、この方法は、容易、迅速など、さまざまな利点がある<sup>11, 12)</sup>。しかしながら、最も重要な欠点は、その質が高くないことである。科学的な妥当性が低いため、PGの目的である医療の改善という目的を達成できない危険性がある。また、根拠や過程などを利用者が検討することができない。

### (2) 公式的合意形成

上記の伝統的な合意形成の欠点を克服するために、合意形成を方法論的に改善するさまざまな試みが行われてきた。代表的な方法を表5に示した<sup>11, 12)</sup>。

第一に、1970年代から現在まで、最も精力的にPG開発に用いられてきたのが、NIHの「合意形成会議」(consensus development conference) である<sup>11)</sup>。この会議は、評価パネルと専門家パネルの2種類から構成されている。まず、評価パネルが、問題の設定に当り、情報の収集と一次的な評価を行い、少数の重要な質問を作成する。この質問と基本的な情報を専門家パネルにあらかじめ配付し、その後、2日半の合意のための会議を開催する。これは、一般公開されており、専門家からの質問の回答と、それに対する評価パネルからの質疑が行われる。最後に、評価パネルが勧告をまとめ、その内容を公表する。すでに、80以上の問題について検討が行われており、この方法は世界各国でも用いられている。この方法では、合意の過程と方法に大きな改善がもたらされたが、合意形成の基準と方法が不明確な点である。

第二に、1980年代の始めから米国医師会で行われてきたのが、診断治療技術評価 (Diagnostic and Therapeutic Assessment) のプログラムである<sup>9)</sup>。これは、必ずしもPGの開発を行うものではないが、医療技術の適切性を評価し伝達するものであり、PGと類似の役割を果たしている。この方法は、1700名

を越える専門家に、検討課題に応じ郵送法による質問票調査を実施するものである。安全性と有効性について、5段階評価を行う。この方法は、専門家間での意見交換や合意形成を行うものではなく、単純な投票により評価を行うものである。

第三は、1980年代半ばから実施されている健康保険（Harvard地域健康計画）の合意形成法である<sup>31</sup>。この最大の特徴は、地域の臨床医の手による臨床アルゴリズムの開発である<sup>32</sup>。その中で、根拠の評価とともに、一定の基準に基づく合意形成が行われている。ただし、いずれも明確な基準が提示されているわけではなく、とくにPGによる医療に質の改善に大きな重点が置かれている。

第四は、1980年代にRANDによって導入された合意形成の公式的な方法である<sup>33</sup>。専門家パネルは、既存の根拠となる情報を提供される。その後、2段階のデルファイ法による合意形成が行われる。会議の前に医療の適切性について9点法で評価を行い、会議ではその情報の集計結果に基づき討議を行う。その後、討議に基づき、適切性の評価を訂正し、再度点数化を行う。この方法は、最近、さまざまな領域や国で、急速に利用されるようになってきている。この方法の問題は、根拠の評価方法などに明確な基準が設定されていないため、その妥当性や質に問題が残されることである。また、合意の結果が、点数化された一覧表の形で提示されるため、その利用には多くの困難が伴う。

### (3) 根拠に基づくガイドライン開発

1980年代に、こうした合意形成の方法と並行して進められてきた別の方法がある。それが根拠に基づくガイドライン開発である（表6参照）<sup>34, 35, 36</sup>。

1979年にカナダ定期健康診断特別委員会（the Canadian Task Force on Periodic Health Examination, CTFPHE）は、その記念碑

的な報告を公表した。そこでは、根拠の質の評価についての明確な基準（5段階）が利用されており、それに基づき利用可能な情報の評価が行われている。その評価結果に基づき、勧告が行われるが、それにも明確な基準（5段階）が用いられている。この方法は、さらに米国予防サービス特別委員会（the US Preventive Services Task Force, USPSTF）に引き継がれ、検診だけでなく一次予防についても系統的な評価が行われている。

また、それと独立に、米国内科学会は、「臨床効果評価プロジェクト」（the Clinical Efficacy Assessment Project）を1980年代に確立し、専門家パネルによる根拠の評価を行い、PGの開発を行っている<sup>37</sup>。その成果は、Ann Intern Med に適宜報告されているが<sup>38, 39</sup>、上記の方法ほど、根拠と勧告との関連は明確ではない。しかしながら、根拠の情報源、評価方法、統合方法など明確な内容を示すとともに、根拠となった情報についても記載されている。

臨床の領域で最も注目を浴びたのは、米国保健医療政策研究局（the Agency for Health Care Policy and Research, AHCPR）である<sup>7, 11, 10</sup>。現在までに19のPGを開発し、プログラムを終了している。この方法は、根拠の明確な評価基準を設定し、それと関連づけて勧告の基準も設定している。その意味では、CTFPHE や USPSTF とほぼ同様な方法論と言ってよい。ただし、それぞれのPGの間には開発方法にはバラツキがあり、統一性に欠けている。なお、このPG開発過程には、公開討議や外部パネルによる評価などが組み込まれており、包括的な検討が保証されている。

上記のように、開発方法によりPGの特徴は大きく異なる。要約すると、表3に示すように、妥当性は、根拠に基づく方法が最も高く、つぎに公式的合意形成、最も低いのが非

公式的合意形成である。こうした妥当性は根拠の統合方法に最も影響されるが、根拠に基づく方法では、系統的な吟味 (systematic review) が明確に行われているため、妥当性が高いと言えよう。ただし、開発のパネルや過程に、使用者が参加する割合は低いため、開発後の利用も必ずしも高いものとは言えない。

## 2) ガイドライン開発方法の内容と過程

根拠に基づくガイドライン開発方法について、その代表例の具体的な内容と過程を以下に要約して示す。

### 1. 米国予防サービス特別委員会

根拠に基づくガイドライン開発のひな形である CTFPHE、とくにそれを継承した USPSTF の開発方法を表7に示した<sup>6, 11)</sup>。

第一に開発の組織である。その要素の一つが、開発の方法論の基礎哲学として、根拠に基づく方法を用いることが設定される。ただし、この方法は妥当性は高いが、その弱点は根拠が極めて限られていることである。つぎの要素が、財政的支援である。PG開発は本格的になればなるほど、人、時間、物、お金が必要となる。そのため、必要な財政的な保証を得る必要がある。つぎに、パネルの構成である。最適の規模と多分野専門家の構成が問題となる。規模は、一次のTFは20名で構成されていたが、二次のTFは10名に縮小されている。専門分野は、第一線の臨床医・臨床疫学者、および非医師専門職に大きく分けられる。通常、混合パネルの方がより保守的であり、かつ信頼性が高いことが指摘されている。また、最適規模は6-10名とされている。最後に課題の選択であるが、デルファイ法を用い、重要疾患のリストを作成し、さらに6つの主観的基準により評点化して優先順位を付ける試みが行われていた。

第二は、根拠の吟味である。まず、効果を

判断するための結果の指標を決定することが必要となる。結果の指標は、通常、最終的な結果と中間的結果の2種類に分けられる。この方法では、最終的結果を用いている。というのも、中間的結果が必ずしも最終的結果につながらないからである。一般的に、最終的結果は、効果評価のアワテ者の誤りを減少させるが、ボンヤリ者の誤りを増加させる危険性が高い。また、最終的結果の情報は極めて限定されている。そこで、逆の危険は高まるが、中間的指標を用いることも、一部に試みられている。そのために、中間的指標と最終的結果との結びつきを推定する因果経路 (causal pathway) の方法が検討されている<sup>17)</sup>。

つぎは、根拠の収集である。この過程は労力と費用が掛かる。通常は、MEDLINE を中心としたコンピュータ文献検索を実施し、手作業検索により補完する。さらに、収集した情報について、その根拠の質を評価する。ここでは、無作為化比較試験 (RCT) を基準とした規則を設定している (5段階評価)。最後に、こうした系統的吟味の結果を記述する。文章により叙述する方法が用いられている。しかし、他の方法では、根拠の一覧表やバランス表を利用している。また、この USPSTF では、メタ・アナリシスや判断分析を独自に実施することはないが、情報としては利用している。

第三は、勧告の作成である。このTFでは、根拠の評価と関連した勧告の基準が設定されている (5段階評価)。この基準を用い勧告を決定するが、その前に、いくつかの問題を検討することが求められる。一つは、根拠の一般集団への外挿 (一般化) 可能性である。根拠の対象となった集団と、一般集団あるいは特殊集団との間で特徴に差が認められるかどうかを検討する。USPSTF は、この適用については保守的な態度を取っている。つぎは、効果の強度である。介入の効果が、統計学的有意差をもって認められたとしても、それが

害を上回るような十分な利益となっているかどうかを検討する。その意味では、利益だけでなく害も評価を行っているが、両者の総合的なバランスの評価については必ずしも定量的な評価を行っていない。判断分析などの方法も利用可能であるが、複雑であり不確実性を除外できないため、組み入れてはいない。もう一つは、費用の問題である。一次のTFでは、費用（費用－効果）の問題については、検討から除外されており、ほとんど臨床的な利益と害に限定されていた。しかしながら、二次のTFでは、根拠の評価に際して、定常的に費用の問題を検討している。

第四は、外部の審査である。USPSTFでは、専門家組織だけでなく、600名を越える専門家（国内外）が勧告案について吟味を行っている。こうした外部審査は、根拠要約の妥当性と完全性を検証し、勧告を改善する上で役立っている。

第五は、他の集団との協力である。USPSTFは、その先行組織であるCTFPHEと方法論を共有するだけでなく、相互の交流会合を開催し、さらに書籍や論文などを協同で出版している。またその他にも、内科学会や家庭医学会との協同作業も実施している。

第六は、影響の評価である。USPSTFの目的はPGの開発であり、その実行結果を評価するものではない。したがって、PG実行の監視は公式に計画されていない。しかしながら、PG開発の目的は、その実行による保健医療の改善にある。その意味では、今後、PG開発後の普及と実行による、保健医療への影響を評価することが重要な課題となる。その指標としては、知識、態度、結果の3種類がある。

第七は、最新化である。USPSTFの最初のPGは1989年に公表された。それ以降、勧告を最新化することが必要となる。というのも、根拠は次々と生みだされているからである。その後、1995年に2版のPGが公表されてい

る。一方、先行するカナダのTFでも、定期的に改定作業を行っている。したがって、PGについては継続的な最新化が必要とされる。

## 2. 保健医療政策研究局

つぎに、医療における根拠に基づくガイドライン開発の代表例である保健医療政策研究局（AHCPR）の例を見てみよう<sup>7, 11, 12</sup>。ただし、内容については、他の組織での開発方法も組み入れたものである。

第一は、課題と過程の定義である。標的疾患の選択を行う必要がある。AHCPRでは、疾患と治療技術について、すでに優先順位付けの方法論と検討を行っている。また、官報により追加課題の提案を促している。つぎに、パネルの構成である。多分野専門家により構成される（臨床医、方法論者、医療経済学者、他の専門職など、15-20名）。これも官報により募集を行うとともに、関連組織などに問い合わせを行う。パネルが組織されると、対象となる介入（医療）を定義する必要がある。そして、患者集団や臨床の場も明確に設定する必要がある。また、根拠の種類として何を利用するか、介入の結果にどのような種類を用いるかも明確にする。

第二は、臨床的利益と害の評価である。PG開発の主な作業は、介入の利益と害である。その情報源は、科学的な根拠と専門家の意見の2つである。前者の科学的根拠の検討では、まず、文献の把握を行う。コンピュータ文献検索を実施し、手作業検索により補完する。つぎに、収集した情報について、その根拠の質を評価し、等級付けを行う。そのための規則を設定しているが（5段階評価）、AHCPRのガイドラインでは統一した規則が用いられていないため、内容にバラツキが認められる。最後に、根拠の統合を行う。ここでは、文章による叙述とともに、根拠の一覧表やバランス表が利用されている。また、メタ・アナリシスや判断分析などを組み入れる方法論の開

発が行われており、そうした作業により、アルゴリズムがまとめられることがある。

後者の専門家の意見は、根拠が不十分な場合に主に検討を行う。内部の専門家パネル（20名程度）、外部の専門家パネル（100-200名程度）から意見を収集するとともに、公開討議などを行う。さらに、必要に応じ、専門家調査も実施する。専門家意見を組み入れる上で、合意形成が行われるならば、そのための方法を選択し、明記する。

第三は、資源と実現可能性の問題の検討である。資源への圧力が高まる中で、医療の資源に関する問題の検討が必要である。それには、費用や、人、技術などが含まれ、PGの成果を保証する意味でも、経済的評価の検討が求められる。また、PGの実現可能性も同様に評価が必要となる。

第四は、勧告の開発である。上記のような利益と害、資源、可能性を検討し、勧告を作成するための系統的な方法を選択する。AHCP Rの勧告では、USPSTFのように、根拠の評価と勧告との関連性が必ずしも明確ではない。しかも、それぞれのPGによって大きな変動をともなっている。また、最終的な勧告については、等級付けの基準を利用している（3ないし5段階）。ただ、これもPGの間で差異が認められる。

第五は、ガイドラインの記述である。AHCP Rでは、3種類のPGが作成されている。一つは医師向けの簡略版であり、もう一つは詳細な情報が記載された本格版である。そして、重要なのは、患者向けのPGである。とくに、患者向けのものは、判断の分かち合い（shared decision making）が十分確立できるように、表現や内容に注意が必要である。とくに、利益と害が、一目で確認できるようなバランス一覧表が有益であることが指摘されている。

第六は、外部審査である。PGについては、科学的な妥当性ととも臨床問題への適切性について、臨床専門家、方法論者、などから

評価を受ける必要がある。また、PGの実行性を検討するためには、関連する専門家、他の組織や行政からも評価を受けるべきである。

### 3. 北部イングランドEBPG開発プロジェクト

英国においても、近年、急速にPGの開発が行われている。その代表的な開発方法の例を表9に示した<sup>19)</sup>。内容と過程は、上記のUSPSTF(CTFPHE)とほぼ同様である。そこで、詳細については省略し、いくつか目立った特徴を記載する。その一つは、パネルのメンバーとして患者が参加していることである。根拠の評価については、EBMの批判的吟味の論文が利用されている。また、根拠の分類、勧告の分類は、CTFPHEのものが準用されている。PGの最新化については3年ごととされている。最後に、PGの実践に際しては、地域的な修飾をしても良いことがしてきされている。

### 4. オンタリオ癌総合センター

カナダはEBM発祥の地であり、伝統的に根拠に基づくPGの開発が盛んである。その中で、地域におけるPG開発例としてこのオンタリオ癌総合センターの例<sup>19)</sup>を、表10に示した。このPG開発は、オンタリオ州政府からの依頼により実施したものである。

開発方法の、1から5までの過程は、今まで検討してきた開発方法と大きな違いはない。その意味では、程度の差はあるものの、根拠に基づくガイドライン開発の方法については、基本的な枠組みが成立しているものと考えられる。この過程の中で、いくつかの注目すべき点について述べる。3のEBR（勧告）の承認は、外部審査とも関連するものであり、実地臨床家の意見を中心として、PGの受容性を検討する。その際、見解の異なる少数意見についても明記する。4のPGの作成においても、柔軟性と信用を得るために、とくに

勧告の修正や修飾を行うべき点について検討を行っている。5の独立した吟味では、作成したPGについて、さらに信用と正統性を高めるために、外部審査が強調されている。

6と7の過程は、とくにPGの実行可能性を調整するものである。6では、臨床医は、その専門責任性として、臨床面からのPGの受容について検討するが、PGを取り巻く非臨床的な環境側面（PGの実行可能性や財政可能性）については、関連者との協同が主な作業となる。7では、PGの発注者に、その採用と実行を計る過程である。それに関連したPGの修正も検討、記録される。

## 5. 米国内科学会系統的総説

近年のPGは、開発から、さらに普及、実行、評価の各段階についての検討に進んでいる。米国内科学会では一連の系統的総説（systematic review）の論文の一つで、系統的総説とPG開発方法との関連について検討を行っている<sup>20)</sup>。図1にその過程を示した。ガイドラインの定式化までは、上記のPG開発と大きな違いはない。一つ特徴的な点は、系統的総説への強調があることである。根拠の統合を行う系統的総説には、質的および量的方法の2つがあるが、PGでは主に質的な統合を行っている。量的方法は、メタ・アナリシスが中心となるが、この点については、従来のPG開発方法では、過程の中に組み込まれていない。その意味では、今後は、こうした方法がPG開発の中で具体的に進められることが予想される。

なお、この開発方法で注目し値するのは、PGをクリニカル・パスウェイに結びつける点である。近年、米国では、経営管理医療（managed care）の下で、医療費を削減し、しかも医療の質を保証する管理が進められている<sup>21)</sup>。臨床現場における具体的な方法の一つが、クリニカル・パスウェイである。この方法は、一定の疾患-治療の組み合わせについ

ては、どのような医療を提供するか、人、資源、時間を標準化させるものである。現在普及しているPGには、この点を明示的に記載したものはまだ認められない。ただし、今後は、こうした領域を意識的に視野に入れたPG開発法が確立されるものと考えられる。

## 3) ガイドライン開発の費用

前にも指摘したが、PG開発に当っては、財政的支援が重要な課題になること。いままでのPG開発にどのような費用を、どの程度要したかについては、ほとんど検討されていない。唯一の例外は、IOM（Institute of Medicine）の報告であり、具体的な数値を挙げて問題の検討を行っている（表11）<sup>2)</sup>。

開発組織は不明であるが、IOMの調査では、一つのPG開発費用は、25万ドルから50万ドルであった。ただし、費用項目については、詳細な記載が無いが、直接費用（スタッフ、パネル人件費など）のみの評価である。関連する調査費用は除外されている。また、NIHの公式的合意形成会議の費用は、8.2万ドルと推定されている。この場合、スタッフの費用、ボランティアの費用は除外されており、かなり過小評価されたものと考えられる。AHCPRのPGでは、初期は20万ドルから25万ドル、後期は35万ドルから80万ドルとなっている。これらは、直接費用のみと考えられるが、詳細は不明である。GAOによる学会のPG開発費用の調査では、5千ドルから13万ドルとなっている。ただし、直接費用のみであり、しかもボランティアの費用は除外されている。後者の費用は、年間25万ドル近くなり、開発に2年を要すると、50万ドルとなる。したがって、実質は、30万ドルから70万ドル近くの費用が掛かっているものと推定される。

以上のように、PG開発には多額の費用が必要となるため、十分な財政的基盤を確立して、計画を立てるようすべきと考えられる。

#### まとめ：ガイドライン開発の課題

P Gは、保健医療のバラツキを減少させ、その適切な利用を促進し、健康改善を確実化するために開発される<sup>1), 2)</sup>。これらの目的を達成するうえで、P Gの長所と短所が、表12に示すようにまとめられている<sup>2)</sup>。

長所としては、多分野の人の参加と、彼らの多様な意見を統合できることが挙げられる。また、P G開発と実践は多くの熱意により支えられている。さらに、根拠の評価とその統合は、信頼の獲得につながる。最後に、多くの利害関係者からの支援と補足が行われる。短所としては、多分野の人の参加により、意見が断片化する危険が生じる。また、P G開発の方法と手技について、質の管理が欠けていることが多い。しかも、P Gの実践にともなう影響について評価が欠けている。

P Gのこうした長所と短所を改善するためには、表13のようなP Gの特性が必要とされることが指摘されている<sup>7)</sup>。第一は妥当性である。P Gは、科学的な根拠に基づくことが求められる。つぎは信頼性である。誰が何処で実施しても、同じようなP Gが再現できるように、明確な過程と構造を設定する必要がある。第三は、臨床適用可能性である。P Gが実際に実行できるような、問題の設定と勧告の内容が必要とされる。第四は、臨床柔軟性である。P Gは、個別の患者や状況に適用できるように、自由度を保証する必要がある。第五は、多分野の参加過程である。保健医療に関係する多様な専門家、あるいは患者が参加することにより、より適切な判断を可能とするようなものでなくてはならない。第六は、再評価の日程である。P Gの内容は、急速に古くなるため、定期的に再評価を行うことができるよう、明確な期限を設定しなければならない。第七は、記録である。方法や実際の過程、根拠などについて、いざでもが再検討できるように、明確で包括的な記載が必要と

なる。

こうした条件を満たすように、さまざまな試みが行われてきたが、その集約とも言えるのが、根拠に基づくガイドラインの開発である。上記で検討してきたように、それにもさまざまな種類があり、それぞれ改善が試みられてきている。また、P Gが実際にこうした条件を満たしているかどうかを検討するために、具体的なチェックリストが、根拠に基づく医療（E B M）グループから提示されている（表14）<sup>21), 22)</sup>。いずれもP G開発方法で重視されていた項目である。

その意味では、わが国でもP Gを開発する上では、根拠に基づくガイドライン開発の方法を用いることが不可欠である。ただし、こうした方法を機械的に適用するのではなく、わが国の状況に適用できるような修飾が必要となる。

#### 文献

- 1) Field MJ, Lohr KN, eds: Clinical practice guidelines, National Academy Press, Washington DC, 1990
- 2) Field MJ, Lohr KN, eds: Guidelines for clinical practice, National Academy Press, Washington DC, 1992
- 3) Woolf SH: Practice guidelines: a new reality in medicine. *J. Recent development, Arch Intern Med*, 150:1811-1818, 1990
- 4) Harris JS: Development, use, and evaluation of clinical practice guidelines. *JOEM*, 39:23-34, 1997
- 5) 久繁哲徳：疾病経営管理，日本製薬工業協会，1998
- 6) Woolf SH, et al: Developing evidence-based clinical practice guidelines. *Ann Rev Public Health*, 17:511-538, 1996
- 7) Agency for Health Care Policy and Research: Report to Congress: HCPR's

- program of patient outcomes research and related activities, AHCPR Pub. No. 95-NO 11, AHCPR, Maryland, 1995
- 8) Canadian Medical Association: Implementing clinical practice guidelines: a handbook for practitioners, 1996
  - 9) Jackson R: Guidelines for clinical guidelines, *BMJ*, 317:427-428, 1998
  - 10) Schott AM, et al: Methods used to develop clinical guidelines in France, *Rev Rhum [Engl Ed]*, 830-836, 1996
  - 11) Woolf SH: Practice guidelines: a new reality in medicine, II. Methods of developing guidelines, *Arch Intern Med*, 152:946-952, 1992
  - 12) Grimshaw J, et al: Developing clinically valid practice guidelines, *J Eval Clin Pract*, 1:37-48, 1995
  - 13) Schoenbaum SC, Gottlieb LK: Algorithm based improvement of clinical quality, *BMJ*, 301:1374-1376, 1990
  - 14) American College of Physicians: Screening for thyroid disease, clinical guideline, part 1, *Ann Intern Med*, 129:141-143, 1998
  - 15) Helfand M, Redfern CC: Screening for thyroid disease, an update, clinical guideline, part 2, *Ann Intern Med*, 129:144-158, 1998
  - 16) Woolf SH: Do clinical practice guidelines define good medical care? *Chest*, 113(3 Suppl): 166s-171s, 1998
  - 17) US Prev Task Force: Screening for adolescent idiopathic scoliosis, *JAMA*, 269:2664-2666, 1993
  - 18) Eccles M, et al: North of England evidence based guidelines development project: methods of guideline development, *BMJ*, 312:760-762, 1996
  - 19) Browman GP, et al: The practice guidelines development cycle, *J Clin Oncol*, 13:502-512, 1995
  - 20) Cook DJ, et al: The relation between systematic reviews and practice guidelines, *Ann Intern Med*, 127:210-217, 1997
  - 21) Hayward RSA, et al: User's guides to the medical literature, VIII. How to use clinical practice guidelines, A. Are the recommendation valid? *JAMA*, 274:570-574, 1995
  - 22) Wilson MC, et al: User's guides to the medical literature, VIII. How to use clinical practice guidelines, B. What are the recommendation and will they help you in caring for your patients? *JAMA*, 274:1630-1632, 1995



表1 診療ガイドラインの背景

1. 臨床医の診療様式と健康サービスの利用における大きなバラツキ
2. 多くのサービスの不適切な利用
3. 多様なサービスの利用と非利用による健康結果改善の不確実性

表2 診療ガイドラインの目的

1. 患者と臨床医の臨床判断の支援
2. 個人と集団の教育
3. 医療の質の評価と保証
4. 保健医療資源の配分の指針
5. 医療過誤の訴訟危険の減少

表3 診療ガイドラインの開発方法

方法	妥当性	根拠の統合	使用者の参加
1. 非公式的合意形成	低	専門家意見	多
2. 公式的合意形成	中	非系統的吟味	中
3. 根拠に基づくガイドライン開発	高	系統的吟味	少

表4 非公式的合意形成の特徴

利点

容易, 安価, 迅速

欠点

質の不良

妥当性の欠如  
 明確な方法の欠如  
 正確な記述の欠如

表5 公式的合意形成の事例

N I H 「合意形成会議」

専門家パネル, 評価パネル  
 検討項目の設定, 情報の収集  
 2.5日の会議, 公開討議,  
 合意は非明示的

米国医師会「診断治療技術評価」

1700名の専門家に郵送質問調査  
 安全性と効果の5段階評価  
 単純な投票, 合意形成は無し

ハーバード地域健康計画

根拠の吟味, 公式的合意形成,  
 臨床アルゴリズム

R A N D 「合意形成」

専門家パネル, 根拠の吟味,  
 2段階デルファイ法(9点),

表6 根拠に基づくガイドライン開発

主な機関

カナダ定期健康診断特別委員会 (CTFPHE)  
米国予防サービス特別委員会 (USPSTF)  
米国内科学会  
米国保健医療政策研究局 (AHCPR)

特徴

根拠の評価基準, 根拠の記述,  
勧告の基準,  
専門家意見の組入れ

表7 米国予防サービス特別委員会の  
ガイドライン開発方法

1. 組織

- 1) 分析哲学 (根拠に基づく接近法)
- 2) 財政支援者
- 3) パネルの構成
- 4) 課題の選択  
デルファイ法  
主観的6基準

2. 根拠の吟味

- 1) 効果判断の結果指標  
中間指標と最終指標
- 2) 因果経路
- 3) 根拠の収集: 文献検索
- 4) 研究の質の評価: 根拠の規則
- 5) 結果の記述  
叙述的総説, 根拠の表,  
バランス・シート,  
メタ・アナリシス,  
判断分析, その他

3. 勧告の作成

- 1) 勧告の規則
- 2) 非明示的問題  
外挿, 効果の強度,  
利益と危険のバランス
- 3) 費用

4. 外部審査

5. 他の集団との協力

6. 影響の評価

- 1) 知識
- 2) 態度
- 3) 結果

7. 最新化

表8 根拠に基づくガイドライン開発方法  
(AHCPRなどの経験から)

- 
1. 課題と過程の定義
    - 1) 標的疾患
    - 2) パネル
    - 3) 介入
    - 4) 患者集団
    - 5) 臨床状況
    - 6) 根拠の種類
    - 7) 結果の指標
  
  2. 臨床的利益と害の評価
    - 1) 科学的根拠の検討
      - 1 文献の把握  
コンピュータ検索,  
手作業検索
      - 2 個別研究の評価  
研究の質の判定と  
等級付けの系統的方法
      - 3 根拠の統合  
叙述的総説, 根拠の表,  
メタ・アナリシス,  
判断分析, その他
    - 2) 専門家意見の考慮
      - 1 意見の収集  
内部の専門家パネル,  
外部の専門家, 公開討議,  
専門家の調査
      - 2 意見の評価方法  
非公式的 (討議, 挙手)  
公式的 (数量的分析)
  
  3. 資源と実現可能性の問題の考慮
    - 1) 資源  
費用, 人, 技術
    - 2) 実現可能性  
臨床医の時間, 訓練, 職員の支援,  
医事法, 診療報酬, 医療政策
  
  4. 勧告の開発
    - 1) 利益・害と資源・可能性を検討
    - 2) 勧告への系統的な接近方法  
公式的な判断規則ないし  
チェックリスト
    - 3) 勧告の等級
  
  5. ガイドラインの記述
    - 1) 簡略版
    - 2) 本格版
    - 3) 消費者版  
判断の分かち合い  
利益と害のバランス一覧表
  
  6. 外部審査
    - 1) ガイドラインの検討  
科学的根拠と臨床問題の検討  
(完全性, バランス, 正確性)
    - 2) ガイドラインの実行性
    - 3) 外部審査のメンバー  
臨床専門家, 方法論専門家
-

表9 イングランド北部の根拠に基づく  
ガイドラインの開発方法

- 
1. ガイドライン開発グループ  
 専門家, 患者, 方法論者
  2. 根拠の吟味と統合
    - 1) 検索戦略  
 MEDLINE
    - 2) 文献の評価  
 EBM作業委員会など
    - 3) 根拠の統合  
 根拠の分類 (CTF)  
 勧告の強度 (CTF)
  3. ガイドラインの外部審査
  4. ガイドライン再吟味の日程  
 3年毎に
  5. 診療での実践  
 地域的な修飾可  
 (根拠と勧告の強度)
- 

表10 診療ガイドラインの開発サイクル  
(カナダ・オンタリオ州総合癌センター)

- 
1. 臨床問題の選択と枠組み  
 課題の優先順位  
 課題の選択  
 結果の選択  
 問題の枠組み
  2. 根拠に基づく勧告 (EBR) の作成  
 既存のガイドラインの検索  
 根拠の収集と統合  
 根拠の強さの等級付け  
 予備的なEBRの作成
  3. EBRの承認  
 EBRの伝達  
 最終EBRへの合意形成と承認  
 少数意見の記載
  4. 診療ガイドラインの作成  
 承認EBRの伝達  
 調整要因の適用  
 臨床診療ガイドラインの作成
  5. 独立した吟味  
 独立した吟味のためにEBR,  
 ガイドラインの提出  
 合意によるガイドラインの調整  
 記録の修正  
 承認のためにガイドラインの提出
  6. 診療政策の交渉  
 ガイドラインの伝達  
 非-臨床家の調整者の採用  
 診療政策の交渉  
 ガイドライン/政策の監視
  7. ガイドラインと政策の採用  
 公式承認のためにガイドライン/政策の提出  
 修正理由の記録
  8. 定期的な最吟味  
 無効期限の設定  
 早期評価を促進する動向の監視
-