

研究結果のまとめ (summary of finding)					
患者数		効果		質	重要性
SSRI		三環系		相対(95%CI)	絶対
5,044	4,510	WMD 0.034 (-0.007 to 0.075)	差異なし	中	重大
1,948/7,032 (28%)	2,072/6,334 (33%)	RRR 13% (5%-20%)	5/100	高	重大
1/100,000/ 治療年	58/100,000/ 治療年	RRR 98% (97%-99%) <sup>s</sup>	6/10,000	中	重大

いエビデンスを基礎としてエビデンス全体の評価を下げてはいけない。推定的な害のもっともらしさは間接的なエビデンスに基づいて判断される。たとえば、動物実験の結果から、ある医薬品の重篤な有害作用について重要な懸念がされる場合には、その有害作用についての人でのエビデンスがどのようなものであっても、全体的なエビデンスの質については

低いグレードがつけられるだろう。推定的な害がありそうなことに関するエビデンスが欠如しているために、介入の正味の(net)利益を評価できないこともある。このような場合には、ガイドライン作成班はさらなる研究の実施を勧める決定を下すであろう。

重大なアウトカムに関するエビデンスのすべてが同一の選択肢を支持しており、そのうちいくつかの

アウトカムについて質の高いエビデンスが得られている場合には、全体的なエビデンスの質は高いといえる。たとえば、心筋梗塞の既往患者において、抗血小板療法が致命的でない脳卒中や心筋梗塞のリスクを減少させるという質の高いエビデンスがある。このとき、すべての原因による死亡率に関するエビデンスの質が中程度のものであり、これが重大なアウトカムと考えられたとしても、全体的なエビデンスの質はなお高いとみなされる。

## お勧め(Recommendations)

その介入による利益は害より多いか

### (Dose the intervention do more good than harm?)

お勧めの作成には、利害得失の考慮が必要である。その際には、個々のアウトカムに対し、黙示的(implicitly)ないし明示的(explicitly)に相対的な価値判断が不可欠である。しかし多くの場合、異なるアウトカムに対する重み付けの判断は困難であり、人によって異なってくる。他の人々を代表して意思決定を行う人々は、影響を受けるユーザーの価値観に関するエビデンスを入手することにより、より確固とした基盤を得ることができる。たとえば、初期乳がん患者を対象とした化学療法についてのお勧めを作成する際には、患者が化学療法の副作用の回避と乳がん再発リスクの低減とを比較した場合、後者にどれだけ相対的重要性をおくかを示すエビデンスを入手できれば、より確かな地位を得ることになる。

コストを考える場合には、まず主たる健康上の利益と害のバランスについて明示的な判断が必要である。その介入による利益は害より多いといえるか。状況や患者群によって利益や害が異なる場合には、お勧めは特定の状況や患者群にのみ適用すべきである。たとえば、心房細動の患者に対し、脳卒中のリスクを低減させるために、出血のリスクが伴うとしてもワーファリンの投与を勧めるべきか考えてみよう。ワーファリンの投与効果を血液検査で定期的にモニタリングできる状況であるか否かで、お勧めやその強さは異なってくる。また、脳卒中のリスクが非常に低い患者(65歳未満の合併症のない患者)とリスクがより高い患者(65歳以上で心不全のある患者)では、絶対的なリスク減少が異なるため、お勧め(な

いしその強さ)が異なる。すなわち、お勧めは特定の患者群や状況に対応したものでなければならない。また、お勧めを作成する際には、不利な(disadvantaged)立場の人々の状況を考慮し、適当な場合には、利用者の状況を考慮して修正を加えることが重要である。

利害得失カテゴリーとして、以下を提案する

---

#### 正味の利益あり(Net benefits)

介入による利益は、明らかに害よりも多い。

#### 利害得失あり(Trade-Offs)

重要な利害得失が存在する。

#### 不確かな利害得失あり(Uncertain trade-offs)

介入による利益が害よりも多いといえるか明らかではない。

#### 正味の利益なし(No net benefits)

介入による利益は、害を上回らないことが明らかである。

---

お勧めを作成する人々は以下の4つの主要因子を考慮すべきである。

- ・ 利害得失。主要なアウトカムに関する推定効果の大きさとその信頼限界、各アウトカムに付与された相対的価値を考慮する。
- ・ エビデンスの質
- ・ 特定の状況を考慮したエビデンスの適用。推定効果の大きさに影響を及ぼす重要な要因(医療機関や医療技術へのアクセスなど)を考慮する。
- ・ 対象集団におけるベースラインリスクに関する不確実性

特定の状況における医療行為へのエビデンスの適用やベースラインリスクに関して不確実性がある場合には、そのお勧めの確信は低くなるであろう。たとえば、ある介入に、重要な利益とともに深刻な有害作用が伴うという場合には、対象集団のベースラインリスクが把握されていない状況では、把握されている場合と比較して、お勧めについてかなり確信が低くなる。

**Box 4 : 価値観とは、正しいか間違っているかの問題ではない**

以下の例は、たとえエビデンスについての合意が得られていても、価値観の違いから、異なる人々はお勧めの内容が異なってくることを示す。

クエスチョン：一般地域住民に対して黒色腫のスクリーニング検査を実施すべきか。

セッティング：米国におけるプライマリ・ケア

ベースラインリスク：一般地域住民(1995年における黒色腫の発症は10万人中13.3人)

参考文献：Helfand et al. *Screening for skin cancer. Systematic evidence review No.2*

Rockville, MD : Agency for Healthcare Research and Quality. April 2001. (AHRQ Publication No 01-S002).

黒色腫のスクリーニング検査の正確度(accuracy)と致命的な黒色腫に関するアウトカムについてのエビデンスの質は非常に低い。スクリーニング検査から生じうる害としては、偽陽性の結果に伴う医療処置などがあるが、これらに関するエビデンスは欠如している。これより、全体的なエビデンスの質は非常に低く、スクリーニング検査から得られる正味の利益は不確かである、と結論付けられる。1件のケースコントロール研究によると、スクリーニング検査を受けていない人に対し、受けている人が致命的な黒色腫を発症するオッズ比は0.37である。また白人男性において、黒色腫が原因で死亡する生涯リスクは0.36%であると推定されている。

以上のエビデンスに基づけば、「スクリーニング検査は実施しない」というお勧めを作成する人が多いであろう。なぜなら、スクリーニング検査から得られる不確かな利益よりも、健康な人にスクリーニング検査を実施した場合に生じうるリスクを回避することの方に重要性を置くからである。しかし一方で、未知ではあるが潜在的な害と比較して、スクリーニング検査から得られるかもしれないわずかではあるが重要な利益に重きを置き、「スクリーニング検査を実施した方がよい」と勧める人もいるかもしれない。こういった状況を鑑み、ガイドライン作成班は、コストについて考慮したうえで、臨床実施へ向けてのお勧めは作成せず、不確実性を減少させて利益得失を明白にするような研究の必要性を提示するだろう。

これは、スクリーニング検査に関するお勧めを考案する際に存在する価値観の違いを示す典型的な例である。急性疾患や慢性疾患の治療に関するお勧めの作成においても同様の問題が持ち上がる。このような場合には、個々の重要なアウトカムの相対的価値や不確実性に照らして、予測される利益と害のバランスについて考慮する必要がある。

お勧め度のカテゴリーとして以下を提案する。

**「お勧め度」のグレード**

しなさい ..... (Do it)	相応の知識のある (well informed) 人々であれば、ほぼ全員 (most) が「実施する」と判断するだろう。
した方がよい ..... (Probably do it)	相応の知識のある人々の大半 (majority) が「実施する」と判断するが、その一方で「実施しない」の判断を下す人々が一部 (substantial minority) 存在する。
しない方がよい ..... (probably don't do it)	相応の知識のある人々の大半が「実施しない」と判断するが、その一方で「実施する」の判断を下す人々が一部存在する。
するな ..... (Don't do it)	相応の知識のある人々のほぼ全員が「実施しない」と判断するだろう。

ある介入について何らかのお勧めが作成されている場合であっても、すべての患者を全く同様に治療しなければならない、ということではない。また、お勧めがあるからといって、意思決定に患者を交えなくてよいということではないし、その他の代替的治療のメリットを説明しなくてもよいということにもならない。しかし、相応の知識のある (well informed) 患者のほぼ全員 (most) が同じ選択をすると考えられるため、代替的治療の相対的メリットについての説明は比較的簡潔でよいだろう。お勧めとは、患者もしくは集団の適切な意思決定を容易にする (facilitate) ためのものである。それゆえ、予測されるアウトカムに関して、エビデンスや個人の価値観や好み (preference) に基づいて人々が選択するであろうと思われる内容が反映されていなければならない。「した方がよい」というお勧めは、その介入の提供に際して、医師は患者の価値観や好みについてより注意深く考える必要があることを示唆している。

利害得失が不明瞭である、あるいはエビデンスに

関する合意が得られていないという理由から、お勧めを提示するのが適切ではない場合もある (Box 4参照)。その理由が質の高いエビデンスの欠如である場合には、お勧めの作成に必要なエビデンスを得るための研究の実施を勧めるべきである。

**健康上得られた利益の増分はコストに見合っているか?  
(Are the incremental health benefits worth the costs?)**

ある介入への支出は、他の介入に支出できる予算の減額を意味するため、それによって得られる健康上の利益の増分 (incremental health benefit) がコストの増分 (incremental cost) に見合った価値があるかという判断 (黙示的であれ明示的であれ) に基づいてお勧めは作成される。コストは利用されたリソースの金銭的価値であり、お勧めにおいて重要な要件であるが、一方で状況に依存し (context specific)、時期によって変化し、実際にどれだけのコストがかかるかを評価するのが困難な場合もある。コストを正確に評価することは容易ではないが、予期される健康上の利益や害と並べて医療上の選択肢の増分コストを明示的に検討することを提案する。関連性があり可能であれば、構成要素のコスト (リソースの使用の違い) についての記載を、重要なアウトカムとともにエビデンス・プロフィールに提示すべきである。リソースの使用の違いに関するエビデンスの質は、前述の、他の重要なアウトカムに関して説明した基準に基づいてグレード化されるべきである。

**本システムは実際にどのように機能するか?  
(How it works in practice)**

表2は、選択的セロトニン再取り込み阻害薬 (SSRI) と三環系抗うつ薬を比較したシステムティック・レビュー (1997年)<sup>40)</sup> から得られたエビデンスに本システムを適用したものである。議論の結果、われわれは、うつ病の重症度と中毒による死亡に関するSSRIと三環系抗うつ薬の相対的効果についての中程度の質のエビデンスと、一過性の副作用に関する質の高いエビデンスが存在する、という点において合意した。つぎに、全体的なエビデンスの質は中程度であり、正味の利益はSSRIの方が優れているとい

**ポイントのまとめ**

- ・エビデンスの質とお勧め度を評価するシステムは、組織により異なっている。
- ・評価システムが異なっていたり、不十分であることは、意思決定プロセスを混乱させ、効率的なコミュニケーションを阻む。
- ・われわれは、エビデンスの質とお勧め度についての判断を下すためのシステムティックで明示的なアプローチを提示した。
- ・このアプローチは、個々のエビデンスの質を判断する際に、研究デザイン、研究の質、一貫性、そして直接性を検討するものである。
- ・お勧め度の判断には、利益と害のバランス、エビデンスの質、適用可能性、およびベースラインリスクの確実性について、すべて考慮されている。

う点でも合意した (うつ病の重症度には差異はなかったが、一過性の副作用と中毒による死亡は三環系抗うつ薬より少なかった)。SSRIに正味の利益が認められるという点では合意したものの、エビデンスの質に関わる不確かさを考慮し、最終的にはSSRIを、「おそらく」 (probably) 使用する、というお勧めとする結果となった。今回、両者の使用コストを比較したエビデンスは入手できなかった。もしコストが検討されていたら、お勧めは違ったものになっていたかもしれない。

**結論 (Conclusions)**

エビデンスの質とお勧め度を評価するシステムにおいては、簡潔性 (simplicity) と明瞭性 (clarity) を両立させる必要がある。システムを単純化すれば判断が黙示的になりやすく、明瞭性を欠くことになる。一方で、明瞭性を高め、意思決定のプロセスをわかりやすくしようとするれば、システムは複雑なものになってしまう。われわれが提唱するシステムでは、簡潔性と明瞭性がうまくバランスするよう試みた。しかし、いかに簡潔で明瞭であったとしても、ユーザーの判断は常に必要である。ここで述べてきたアプローチは、構造化された熟考 (structured reflection) のための枠組みを提供するものであり、適切な判断を確実にを行う助けとなるが、ユーザー自身の判断の必要性をなくしてしまうものではない。

文献 (References)

- 1) Canadian Task Force on the Periodic Health Examination. The periodic health examination. *CMAJ* 1979;121: 1193-254.
- 2) Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1986;89(suppl 2): 2-3S.
- 3) Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Arch Intern Med* 1986;146: 464-5.
- 4) Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1989;95: 2-4S.
- 5) Cook DJ, Guyatt GH, Laupacis A, Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. Antithrombotic therapy consensus conference. *Chest* 1992;102(suppl 4): 305-11S.
- 6) US Department of Health and Human Services, Public Health Service, Agency Health Care Policy and Research. Acute pain management: operative or medical procedures and trauma. Rockville, MD: Agency for Health Care Policy and Research Publications, 1992. (AHCPR Pub 92-0038.)
- 7) Gyorkos TW, Tannenbaum TN, Abrahamowicz M, Oxman AD, Scott EA, Millson ME, et al. An approach to the development of practice guidelines for community health interventions. *Can J Public Health* 1994;85(suppl 1): S8-13.
- 8) Hadorn DC, Baker D. Development of the AHCPR-sponsored heart failure guideline: methodologic and procedural issues. *Jt Comm J Qual Improv* 1994;20: 539-54.
- 9) Cook DJ, Guyatt GH, Laupacis A, Sackett DL, Goldberg RJ. Clinical recommendations using levels of evidence for antithrombotic agents. *Chest* 1995;108(suppl 4): 227-30S.
- 10) Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ, et al. Users' guide to the medical literature IX: a method for grading health care recommendations. *JAMA* 1995;274: 1800-4.
- 11) Scottish Intercollegiate Guidelines Network (SIGN). Forming guideline recommendations. In: A guideline developers' handbook. Edinburgh: SIGN, 2001. (Publication No 50.) [www.sign.ac.uk/guidelines/fulltext/50/section6.html](http://www.sign.ac.uk/guidelines/fulltext/50/section6.html) (accessed 8 Feb 2004).
- 12) US Preventive Services Task Force. Guide to clinical preventive services. 2nd ed. Baltimore: Williams and Wilkins, 1996: xxxix-lv.
- 13) Eccles M, Clapp Z, Grimshaw J, Adams PC, Higgins B, Purves I, et al. North of England evidence based guidelines development project: methods of guideline development. *BMJ* 1996;312: 760-2.
- 14) Centro per la Valutazione della Efficacia della Assistenza Sanitaria (CeVEAS). Schema di grading CeVEAS. <http://web1.satcom.it/interage/ceveas/html/doc/45/GLICO.pdf> (accessed 18 May 2004).
- 15) Guyatt G, Schunemann H, Cook D, Jaeschke R, Pauker S, Bucher H. Grades of recommendation for antithrombotic agents. *Chest* 2001;119:3S-7S. [www.chestjournal.org/content/vol119/1\\_suppl/](http://www.chestjournal.org/content/vol119/1_suppl/) (accessed 8 Feb 2004).
- 16) Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B, Dawes M. Levels of evidence and grades of recommendations. Oxford: Oxford Centre for Evidence-Based Medicine. [www.cebm.net/levels\\_of\\_evidence.asp](http://www.cebm.net/levels_of_evidence.asp) (accessed 8 Feb 2004).
- 17) National Health and Medical Research Council. How to use the evidence: assessment and application of scientific evidence. Canberra: AusInfo, 2000. [www.health.gov.au/nhmrc/publications/pdf/cp69.pdf](http://www.health.gov.au/nhmrc/publications/pdf/cp69.pdf) (accessed 8 Feb 2004).
- 18) Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ* 2001;323: 334-6.
- 19) Roman SH, Silberzweig SB, Siu AL. Grading the evidence for diabetes performance measures. *Eff Clin Pract* 2000;3: 85-91.
- 20) Woloshin S. Arguing about grades. *Eff Clin Pract* 2000;3: 94-5.
- 21) Guyatt GH, Schunemann H, Cook D, Pauker S, Sinclair J, Bucher H, et al. Grades of recommendation for antithrombotic agents. *Chest* 2001;119: 3-7S.
- 22) Atkins D, Best D, Shapiro EN, eds. Third US Preventive Services Task Force: background, methods and first recommendations. *Am J Prev Med* 2001;20: 3(suppl):1- 108.
- 23) Woolf SH, Atkins D. The evolving role of prevention in health care: contributions of the US Preventive Services Task Force. *Am J Prev Med* 2001;20: 3(suppl):13-20.
- 24) Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, et al. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med* 2001;20:3 (suppl): 21-35.
- 25) Briss PA, Zaza S, Pappaioanou M, Fielding J, Wright-De Aguiro L, et al. Developing an evidence-based guide to community preventive services-methods. *Am J Prev Med* 2000;18 (suppl 1): 35-43.
- 26) Zaza S, Wright-De A, Briss PA, Truman BI, Hopkins DP, Hennessy MH, et al. Data collection instrument and procedure for systematic reviews in the guide to community preventive services. *Am J Prev Med* 2000;18(suppl 1): 44-74.
- 27) Greer N, Mosser G, Logan G, Halaas GW. A practical approach to evidence grading. *Jt Comm J Qual Improv* 2000;26: 700-12.
- 28) West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, et al. Systems to rate the strength of scientific evidence. Rockville, MD: Agency for Healthcare Research and Quality, 2002: 64-88. (AHRQ publication No 02-E016.)
- 29) Guyatt G, Drummond R, eds. Users' guide to the medical literature. Chicago, IL: AMA Press, 2002: 55-154.
- 30) Clarke M, Oxman AD, eds. Assessment of study quality. Cochrane reviewers' handbook 4.1.5 section 6. In: Cochrane Library. Issue 4. Oxford: Update Software, 2002.
- 31) Juni P, Altman DG, Egger M. Assessing the quality of randomised controlled trials. In: Egger M, Davey Smith G, Altman DG, eds. Systematic reviews in health care: meta-analysis in context. London: BMJ Books, 2001: 87-121.
- 32) West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, et al. Systems to rate the strength of scientific evidence. Rockville, MD: Agency for Healthcare Research and Quality, 2002: 51-63. (AHRQ publication No 02-E016.)
- 33) Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in health-care trials (Cochrane methodology review). In: Cochrane Library Issue 4. Oxford: Update Software, 2002.
- 34) Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 2001;286: 821-30.
- 35) Kleijnen J, Gotzsche P, Kunz RA, Oxman AD, Chalmers I. So

- what's so special about randomisation? In: Chalmers I, Maynard A, eds. *Non-random reflections on health care research: on the 25th anniversary of Archie Cochrane's effectiveness and efficiency*. London: BMJ, 1997: 93-106.
- 36) Lacchetti C, Guyatt G. Surprising results of randomized controlled trials. In: Guyatt G, Drummond R, eds. *Users' guide to the medical literature*. Chicago, IL: AMA Press, 2002: 247-65.
- 37) Hulley S, Grady D, Bush T, Furberg C, Herrington D, Riggs B, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. *JAMA* 1998;280: 605-13.
- 38) Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. Principal results from the women's health initiative randomized controlled trial. *JAMA* 2002;288: 321-33.
- 39) Dans A, McAlister F, Dans L, Richardson WS, Straus S, Guyatt G. Applying results in individual patients. In: Guyatt G, Drummond R, eds. *Users' guide to the medical literature*. Chicago, IL: AMA Press, 2002: 369-84.
- 40) McAlister F, Laupacis A, Wells G. Drug class effects. In: Guyatt G, Drummond R, eds. *Users' guide to the medical literature*. Chicago, IL: AMA Press, 2002: 415-31.
- 41) Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The cardiac arrhythmia suppression trial. *N Engl J Med* 1991;324: 781-8.
- 42) Riggs BL, Hodgson SF, O'Fallon WM, Chao EY, Wahner HW, Muhs JM, et al. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med* 1990;322: 802-9.
- 43) Worster A, Preyra I, Weaver B, Haines T. The accuracy of noncontrast helical computed tomography versus intravenous pyelography in the diagnosis of suspected acute urolithiasis: a meta-analysis. *Ann Emerg Med* 2002;40: 280-6.
- 44) Worster A, Haines T. Does replacing intravenous pyelography with noncontrast helical computed tomography benefit patients with suspected acute urolithiasis? *Can Assoc Radiol J* 2002;53: 241.
- 45) Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: evidence from published meta-analyses. *BMJ* 2003;326: 472.
- 46) Bross IDJ. Pertinency of an extraneous variable. *J Chron Dis* 1967;20: 487-95.
- 47) Thompson DC, Rivara FP, Thompson R. Helmets for preventing head and facial injuries in bicyclists. *Cochrane Database Syst Rev* 2000;(2): CD001855.
- 48) Devereaux PJ, Choi PT, Lacchetti C, Weaver B, Schunemann HJ, Haines T, et al. A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals. *CMAJ* 2002;166: 1399-406.
- 49) North of England Evidence Based Guideline Development Project. Evidence based clinical practice guideline: the choice of antidepressants for depression in primary care. Newcastle upon Tyne: Centre for Health Services Research, 1997.

## 診療ガイドライン作成・活用ガイド\*

中山 健夫\*\*

〔要旨〕 根拠に基づく医療 (EBM) の手法を用いた診療ガイドラインに対する関心が、医療者のみならず一般の人々のあいだでも高まっている。診療ガイドラインの定義は「特定の臨床状況のもとで、医療者と患者の意思決定を支援する目的で、系統的に作成された文書」であるが、その意義は国内では必ずしも適切に理解されていない。本レビューでは、国内における診療ガイドラインの作成とその活用に関する課題と展望の紹介を行いたい。

## はじめに

診療ガイドラインは、米国の Institute of Medicine により「特定の臨床状況のもとで、医療者と患者の意思決定を支援する目的で、系統的に作成された文書」と定義されている<sup>1)</sup>。また英国・スコットランドで診療ガイドライン作成・普及の責任を担っている SIGN (The Scottish Intercollegiate Guidelines Network) は、「私たちの目的は、現在のエビデンスに基づく治療の有効性に関する推奨を含む国レベルの診療ガイドラインの作成、普及を通じて、診療とアウトカムのばらつきを減じ、スコットランドにおける患者ケアの質を向上させることである」として、診療ガイ

ドラインに求められる特性とその役割を簡潔に表現している。

国内では厚生省 (当時) は 1999 年度の厚生科学研究から、EBM の手法を用いた診療ガイドライン作成を開始した (図 1)。それまでは診療ガイドラインの作成において、従来は根拠とする文献の選択・入手法、評価法、勧告 (推奨) の決定法などが明示されず、何人かの主導的立場にある臨床医のコンセンサスによってつくられることが一般的であった。厚生 (労働) 科学研究によるプロジェクトでは、臨床疫学者や生物統計学者のような研究方法論の専門家や、適切な文献を検索するために医学図書館員の参加が強く求められた<sup>2)</sup>。現在は公的研究費の枠外でも多くの学会が独自に診療ガイドライン作成に着手しており、今後、その方向が拡大していくことが予想されている。

医療者にとってもっとも伝統的な情報源である教科書は、はじめて医学に接する者が解剖学、病態生理学、または記述的な疫学の知識、疾病単位 (disease entity)、疾病相互の関係などの専門的

キーワード：診療ガイドライン、根拠に基づく医療、推奨度、利害関係者、医師の裁量

\* Guides to development and utilize clinical practice guidelines

\*\* T. Nakayama (助教授)：京都大学大学院社会健康医学系専攻健康情報学分野。

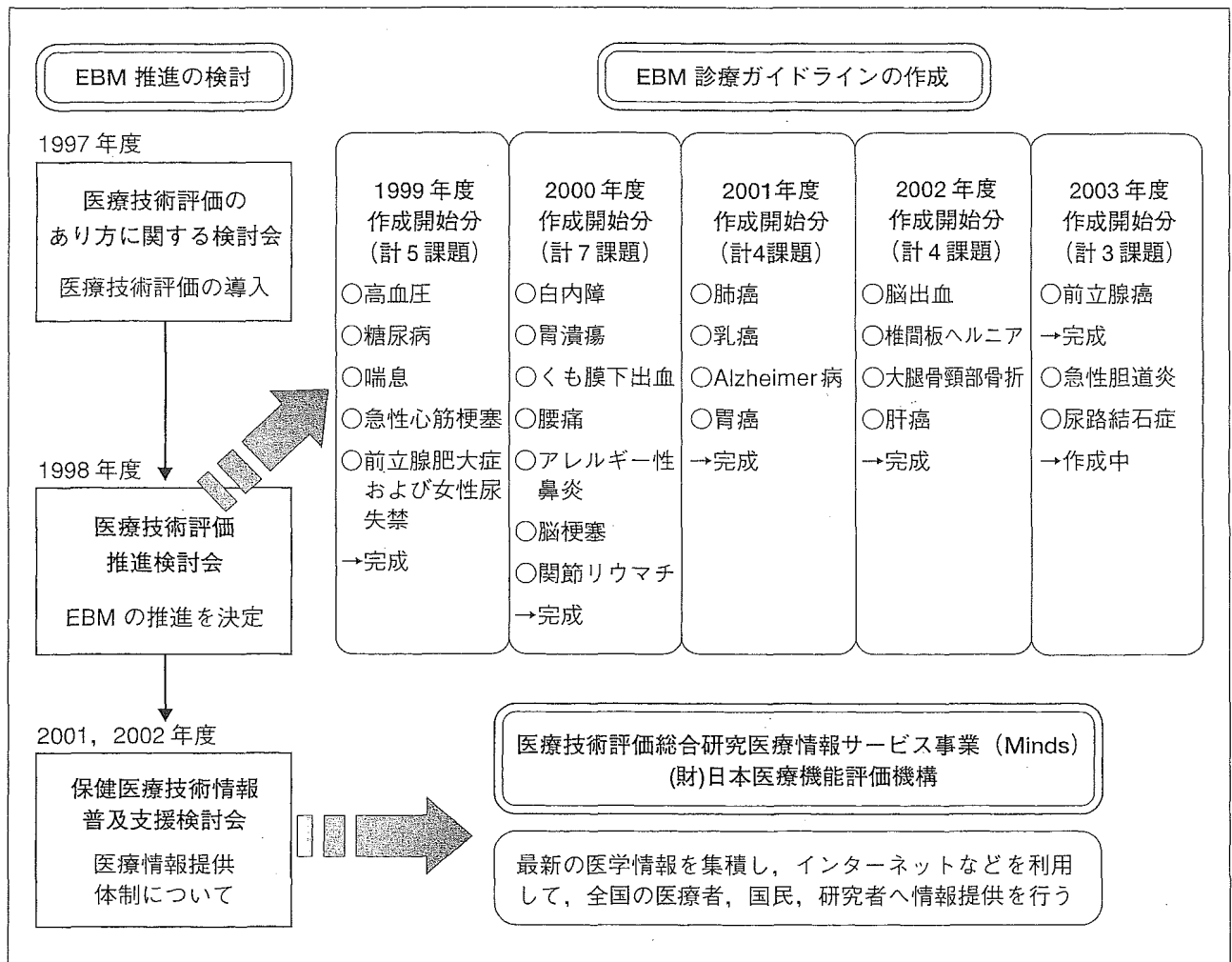


図1. EBMに対する厚生労働省の取り組み

知識の枠組みを学ぶものであり、専門家はそのバックボーンを形成するのに不可欠の知識体系を示すものといえる。そのような教科書的記述が、臨床現場におけるリアルタイムの問題解決には必ずしも直結しないのに対して、診療ガイドラインは臨床現場で向き合う特定の問題に対して一定の方向性(推奨)を示し、直接的に現場の判断を支援することを目指す。現在、確立しつつあるEBMの手法に基づくガイドラインでは、できる限り客観的なエビデンスに基づいて推奨を行うことが求められており、一定の明示的な作成プロセスを踏むことが重視されている。

### I. 診療ガイドラインの構造と作成手順

ガイドラインを作成するテーマを設定したら、

その問題の関係者を集めて作成グループを組織する。日本では学会の主導でガイドライン作成がすすめられることが多く、学会関係者という、ある意味では「仲間」、わるくいえば「内輪」で作業がすすめられることが多いが、欧米では stakeholders(利害関係者)を広く集め、さまざまな意見対立の中から合意を形成するプロセスが重視される。日本でも今後はガイドライン作成、とくに医療の領域でも社会性の高いテーマについては、このような取り組みの重要性が増していくであろう。

EBMの手法による診療ガイドライン作成は、まず推奨を示すべき臨床的課題(clinical questions)をあげ、それに対してシステマティック・レビューの方法に準じて、関連文献の系統的検索



と吟味を行う<sup>3)</sup>。臨床的課題は、EBMの第1ステップである「疑問の定式化」の方法を参考に、「どんな患者・疾患・病態に対し(patient), 何を行い(intervention), どんな臨床的効果を期待するのか(outcome)」という視点で作成する。これに「何と比較して(comparison)」を加え、EBMでは頭文字から“PICO”と称される。

推奨度の決定は診療ガイドラインの核心である。あるテーマの文献が収集され、その内容が抽出された「診療エビデンス(clinical evidence)集」は、方向性を示す推奨を含まない点で診療ガイドラインと決定的に異なる。推奨度は得られたエビデンスの質(レベル)に大きく影響されるが、それだけで決定されるものではない。推奨度は、表1の要素を勘案して総合的に判断されるものである<sup>4)</sup>。

推奨度の表示形式を表2に示す。国内では1990年代前半に米国AHCPR [Agency for Health Care Policy and Research(現 Agency for Healthcare Research and Quality: AHRQ)] が提案した方式が多くガイドラインで用いられているが、判定が困難な場合も少なくない。とくに「推奨度C: 根拠がないので推奨できない」の解釈については混乱が多いため、今後はこのAHCPRの4段階方式ではない表示方法を考慮することが望ましい。

日本脳卒中学会の合同ガイドライン(2001年)ではC判定を「C1: 行うことを考慮してもよいが、十分な科学的根拠はない」と「C2: 科学的根拠がないのですすすめられない」に区別している。棚橋は、同ガイドラインにおける急性期治療の推奨71のうち、Aが0件、Bが6件、C1が53件、C2が8件、Dが4件であり、C1が75%を占めていることを報告している<sup>5)</sup>。これは、必ずしも「日本の臨床医学は、高いレベルのエビデンスに基づいて行われていない」という否定的な意味ではなく、「実際に臨床で行われていることの多くは高いレベルのエビデンスで裏付けられていないものである」という臨床の現実を意味していることとらえることもできる。

近年の諸外国での推奨度の表現方法としては次

表1. 推奨度の決め方  
以下の要素を勘案して総合的に判断する。

- |                                                                                                                                                                                                                                             |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ol style="list-style-type: none"> <li>1) エビデンスのレベル</li> <li>2) エビデンスの数と結論のばらつき<br/>(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勧告は強いものとなる。必要に応じてメタアナリシスを行う)</li> <li>3) 臨床的有効性の大きさ</li> <li>4) 臨床上の適用性</li> <li>5) 害やコストに関するエビデンス</li> </ol> |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

のような提案がある。臨床における予防医学的テーマの包括的ガイドラインである、米国の『Preventive Services Task Force(第3版)』では、ABCDの4段階の推奨度に新たにI判定を加えている。推奨度C, Iとも明確な推奨がむずかしい課題であるが、Cは複数の信頼しうる研究報告が存在するが、利益と害が拮抗しているため一定の推奨が困難であるのに対し、Iはエビデンスがないか、その質が低いために推奨が決定できないケースがあたるとしている。表2の4に示すようにカナダの予防医学に関するタスクフォースも米国に類似した、しかし独自の推奨度分類を提案している。これは医療の非専門家も含め関係者の誤解・齟齬を減らす意味で有益な視点と思われる。同タスクフォースは、推奨度Cのようにエビデンスが明確でない場合の意思決定・行動の基準を示している(表3)。今後の国内におけるガイドライン作成においても検討をすすめるべき問題の一つであろう。

## II. 診療ガイドラインと「医師の裁量」

診療ガイドラインは日常診療の対象となる多くの症例に対して参考になるが、診療ガイドラインを個々の患者の状況、特性に応じて柔軟に利用することが必要である。欧米で「診療ガイドライン」は“starting point for discussion”として当事者の意思決定を「支援する」役割を担う。しかし今後、医療者の想定を越える領域でも、医療訴訟の判断基準として「診療ガイドライン」が社会的に

表2. 推奨度の形式

- 1) 米国 Agency for Health Care Policy and Research(1994年)
  - A: 行うよう強くすすめられる
  - B: 行うようすすめられる
  - C: 行うようすすめるだけの根拠が明確でない
  - D: 行わないようすすめられる
- 2) 脳卒中合同ガイドライン委員会(2004年)
  - A: 行うよう強くすすめられる(少なくとも一つのレベルIの結果)
  - B: 行うようすすめられる(少なくとも一つのレベルIIの結果)
  - C1: 行うことを考慮してもよいが, 十分な科学的根拠はない
  - C2: 科学的根拠がないので, すすめられない
  - D: 行わないようすすめられる
- 3) 米国臨床予防医学タスクフォース(2001年)
  - A: 臨床医は, 対象となる患者に定期的に介入(ここでは予防サービス, 以下介入とする)を提供することを強く推奨する(USPSTFは, 介入が重要な健康アウトカムを改善するとする強い科学的根拠を得て, 利益は害よりはるかに大きいとの結論を下している)
  - B: 臨床医は定期的に介入を対象となっている患者に提供しよう推奨する(USPSTFは少なくとも, 介入が重要な健康アウトカムを改善するとする十分な科学的根拠を得て, 利益は害より大きいとの結論を下している)
  - C: 定期的な介入の提供についての是非に対する推奨は行わない(USPSTFは, 少なくとも介入が健康アウトカムを改善するとするかなりの科学的根拠を得たが, 利益と害のバランスが, 一般的な勧告を正当化するにはあまりにも接近しているとの結論を下している)
  - D: 介入を定期的に非症候患者に提供しないよう推奨する(USPSTFは少なくとも, 介入が効果的でない, あるいは害が利益より大きいとするかなりの科学的根拠を得た)
  - I: 介入を定期的に提供することの是非についての科学的根拠が不十分であるとする結論を下す(介入が効果的であるとする科学的根拠の欠如, その質の低さや矛盾, また利益と害のバランスが決定できない)
- 4) カナダ・予防医学に関するタスクフォース(2003年)
  - A: 予防的介入を推奨しうる十分な(good)エビデンスがある
  - B: 予防的介入を推奨しうる適正な(fair)エビデンスがある
  - C: 既存のエビデンスは一致しておらず, 予防的介入を肯定・否定する推奨は決定できない. しかし, 他の要因が意思決定に影響しうるであろう
  - D: 予防的介入を行わないことを推奨する適正な(fair)エビデンスがある
  - E: 予防的介入を行わないことを推奨する十分な(good)エビデンスがある
  - I: 推奨を決定するにはエビデンスが量的・質的に不十分である. しかし, 他の要因が意思決定に影響しうるであろう

認識されていくことの可能性は否定できず, それを確固とした判断の指針と捉える医療消費者も増えていくことが予想される. そうなれば臨床医が過剰な“defensive medicien”への傾向を強めることも懸念される. 「診療ガイドライン」の意義と適切な社会的位置づけは, 医師だけではなく法律家や一般の人々とも協力して取り組むべき重要課題である.

診療ガイドラインの意義が期待とともに語られるのと同程度に, 診療ガイドラインに対する不安, 疑問も投げかけられている. その一つがガイドラインという存在のもつ拘束力である. ガイドライン(guideline)は指針ともいわれるが, この拘束力については国内では十分な議論がされていない. 『疫学辞典(第4版)』によると<sup>6,7)</sup>, 関連する用語として, 指令(directive), 勧告または推奨

(recommendation)があげられ、「指令は勧告(推奨)よりも強く、「勧告」(推奨)は指針よりも強い、北米では指針と勧告(推奨)は同等である」と述べられている。この定義からも明らかなように、ガイドラインの役割はあくまで意思決定の支援であり、現場における個別の臨床行為・意思決定を規制するものではない。参考までに指令(direct)よりも拘束力が強いのは規制(regulation)であり、この場合はそれに従わないと罰則を受けるなど、直接的な不利益が生じることが多い。Recommendationは従来「勧告」と訳されることが多かったが、「勧告」は医療法にもみられる法律用語であり、暗黙の拘束力が意図されている。一般の人々や法律家を交えたガイドラインに関する社会的議論に向けては、誤解を避けるために「勧告」ではなく「推奨」を用いることがより適切であろう。

診療ガイドラインがカバーするケースについて、Eddyは60～95%の患者にとどまると述べ、95%以上の患者に適応される「スタンダード」、反対に50%ほどの患者にしか適応されないものは「オプション」と区別している<sup>9)</sup>。しかし同じ「診療ガイドライン」でも、慢性疾患で予後への影響要因が多様なものと、進行癌に代表されるように特定の治療が予後に強く影響するものと、科学的な意味での拘束力は異なるであろう。したがって、診療ガイドラインに関する議論は総論とともに、特定の臨床状況における各論的な検討を並行して深めていく必要がある。その意味でさまざまな状況で生じる診療ガイドラインをめぐる関係者の齟齬や葛藤、それらへの対応の事例を蓄積していくことは、診療ガイドラインの適切な利用・普及の視点からおおいに重要といえる。

診療ガイドラインは臨床現場に大きな影響を与えるものであるが、臨床医の「プロフェッショナル・フリーダム」との整合が問題とされることが多い。長沢・津谷は日本語の「プロフェッショナル・フリーダム」と英語の“professional autonomy”の語義を、歴史的、文献計量学的に明らかにすることによって、「自律」をベースとする「積極的自由」は職業団体によって重視されるべ

表3. エビデンスが明確でない場合の意思決定の基準(カナダ・タスクフォース)

- |                                                                                                                                                                                                                                                              |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>・意思決定における患者自身の参加を促すこと</li> <li>・害を最小化すること</li> <li>・強い必要性が明らかな場合に関してのみ、大きな変化を主張すること</li> <li>・不要な「ラベリング」を避けること</li> <li>・益の不確かな高価な手技を避けること</li> <li>・疾病負担が大きい状況に焦点を当てること</li> <li>・ハイリスクグループの特別のニーズに配慮すること</li> </ul> |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

きもので、日本語でも「プロフェッショナル・オートノミー」と称されることが適切であるとしている<sup>9)</sup>。この提案によれば医師はプロフェッションとして、積極的に診療ガイドラインの作成や、改訂のための意見表明を行うことが期待されることを意味し、「診療ガイドラインは医師の裁量を拘束する」という従来からの危惧に対して、新しい視点を示すともいえる。

### III. 診療ガイドラインの評価

有効な診療ガイドラインの要件として、Grimshawは①内容が妥当か？、②普及し利用可能か？、③実際に利用されているか？、④患者アウトカムの改善をあげ、診療ガイドラインの臨床行為への影響・患者アウトカムの改善の視点からレビューを行った<sup>10)</sup>。その結果、59の診療ガイドラインのうち55で臨床行為の改善がみられたこと、患者アウトカムへの影響を検討した11論文では2論文を除いてアウトカムが改善したことを報告している。小山らは1991～2000年にかけて発表された、学会または国家機関レベルで作成された治療に関する診療ガイドラインが患者アウトカムに与えた効果を、ランダム化比較試験または準実験的研究デザインにより検討した研究のシステマティック・レビューを行った<sup>11)</sup>。17篇の論文が同定され、うち10篇(59%)で患者アウトカムの改善が示されていた。また、患者アウトカムの改善が示されなかった研究のうち、2篇ではガイドライン遵守率が低く、3篇ではガイドライ

ン遵守率が報告されていなかった。本研究の結果は、診療ガイドラインを診療上用いることにより患者アウトカムが改善される可能性を強く示唆している。

診療ガイドラインの評価に関する国際的な取り組みとして、European Union(EU)を中心に発足したAGREE共同計画(Appraisal of Guidelines Research and Evaluation)がある。AGREEのグループは発表された診療ガイドラインの質を調べるため、次に示す6領域23項目と総合評価からなる評価手法を提示している<sup>12,13)</sup>。

1) 範囲と目的(項目1~3): ガイドラインの全体的な目的, 具体的な臨床問題と対象とする患者集団に関する項目

2) 利害関係者の関与(項目4~7): ガイドラインがそのユーザーとして意図された者の見方をどれほど代表するものであるかに関する項目

3) 開発の厳密さ(項目8~14): 根拠を集め集約するのに用いられた過程と, 推奨を導き出す方法, その更新に関する項目

4) 明快さと提示(項目15~18): ガイドラインの言葉と形式に関する項目

5) 適用性(項目19~21): 考えられるガイドライン適用, 組織的, 行動的, 経済的影響に関する項目

6) 編集の独立性(項目22~23): ガイドライン作成グループが利害の衝突を生じる影響力から独立していることに関する項目

内容をみればわかるようにAGREEによる診療ガイドラインの評価は臨床医の行動や患者アウトカムの変化ではなく, 診療ガイドライン作成の「枠組み」の評価に焦点を当てている。2003年9月には米国が中心となって行われたThe Conference on Guideline Standardization(COGS)の成果が報告された<sup>14)</sup>。COGSの提案は作成された診療ガイドラインを報告するさいのチェック項目であり, すでに発表された診療ガイドラインの評価法としてのAGREEの提案と表裏一体である<sup>15)</sup>。COGSの提案は推奨が導き出された過程の透明性を確保するだけでなく, 作成者, 目的, 資金提供者に始まり, 可能な場合は有効期

限を示すことを求めるなど, その診療ガイドラインのつくられた枠組みに関する情報を重視している。これは情報学でいうメタデータ, すなわち「データに関するデータ」の一種であり, 個々の情報のもつ意味を利用者が判断するさいの有益な手がかりとなる。このような国際的な標準様式に沿ったかたちで診療ガイドラインを作成し報告することが, 今後, 質の高い診療ガイドラインを世に出すための必須要件となっていくことが予想される。

#### IV. 医学研究・診療ガイドライン作成における「利害の衝突」

「利害の衝突(conflict of interest)」または「相反利益(competing interest)」は, 海外ではすでに学術活動の公正さに対する大きな脅威として認識されており, 国内でもより多くの関心を向けられるべき問題である。医学雑誌編集者委員会(International Committee of Medical Journal Editors: ICMJE)〈<http://www.icmje.org/>〉による「生物医学雑誌の統一投稿規程」では, 「著者, 査読者, 及び編集者といったピア・レビュー(査読)及び出版過程に関わる関係者が, 自らの判断に不適切な影響を与え得る活動に関わっている場合, それが現実に影響を与えるかどうかは別として, 寄稿された原稿に関する『利害の衝突』が存在する」としている。企業との財政上の関係(たとえば雇用, コンサルタント業, 株式の所有, 謝礼, 専門家の立場での証言などを通じた関係)は, 直接的なものでも直系の家族を通じたものでも, これは通常もっとも重大な「利害の衝突」としてみなされる。それ以外に個人的人間関係, 学術的な競合, および知的情熱なども原因となる。同規定では査読および出版にかかわる関係者らは自らの利害衝突の状況について明らかにし, 他の関係者が, それによって自らが被るかもしれない影響を自分で判断できるように, そういった情報がアクセス可能することを求めている。

原著論文における「利害の衝突」の影響を知るのに比べ, 総説論文や論評では「叙述のニュアンス」に対する影響となるため, その察知はより困

難である。個々の研究論文の研究デザインや、ある臨床的疑問に関する利用可能なエビデンスに対するエビデンス・レベルの判定はある程度、客観的に可能であるが、行動の方向性を示す推奨度の決定には多くの要因が影響し、作成者の恣意がはいりやすい。多くの論文情報が集約され、推奨のかたちで方向性が示される診療ガイドラインでは、「利害の衝突」が無視できない影響を与えることに、関係者はより多くの注意を向ける必要がある。Lenzerはガイドライン上でのtPAの「格上げ」に米国心臓学会や医師への企業献金が影響していた可能性を指摘し<sup>16)</sup>、またChoudhryらはガイドライン作成者の58%は研究資金供出を受けていたことを報告している<sup>17)</sup>。これまで国内で作成されてきた診療ガイドラインでは、作成者の「利害の衝突」の開示はほとんど行われていない。ガイドライン作成の途中、または発表後に企業が直接、間接に自社製品の記述に対して影響を与えようとした事例が聞かれることも事実であり、今後、国内でも作成プロセスのいっそうの透明性向上が強く望まれる。現在作成中のいくつかのガイドラインでは、その記載が加えられつつあり(例：日本褥瘡学会による褥瘡管理ガイドライン、日本耳科学会による急性中耳炎ガイドラインなど)、この方向性が広く認知されていくことが望まれる。

## V. 診療ガイドライン作成における「患者参加」について

患者と医師の関係からは、患者・家族と医療者の対話の結節点としての診療ガイドラインが機能する可能性を探る意義は大きい。前述のAGREEによる評価項目の一つとして、「患者の視点や選好は考慮されたかどうか」が問われており、「診療ガイドライン開発にあたって、患者の経験と期待に関する情報を知っておかねばならない。そのための方法として開発グループに患者の代表を含める、患者のインタビューから情報を得る、開発グループが患者の経験に関する文献をレビューするなどである。この手順が行われたという記述がなければならない」ことが求められている。

患者の代表者が診療ガイドライン作成に関与することは、厚生労働科学研究による診療ガイドライン作成の方針を示した「診療ガイドラインの作成の手順 ver.4.3」<sup>4)</sup>でも述べられているが、国内ではその取り組みは遅れている。しかし欧米では、診療ガイドラインを作成するのは「その領域の権威」に限定されず、対象課題が設定されると、それに応じたさまざまな“stakeholders(利益関係者)”がそれぞれの立場で発言の機会を与えられることが一般的である。診療ガイドライン作成の初期から「重要な関係者」である患者の参加を得ることは決して容易ではないが、現場で臨床医と患者の双方の視点で有用な診療ガイドラインが作られていくためには、医療者が向き合うべき重要な課題である<sup>18)</sup>。

実際に患者の視点を取り入れるためには、準備期間が必要であろう。今のままで何人かの患者がその代表としてガイドライン作成の会議に参加したとしても、実質的な議論への十分な貢献は期待できない。それは患者にその力がないのではなく、「必要な知識を事前に習得する機会が提供されていないこと」と「個人への依存度が高く、システム化されていないこと」が大きな原因と思われる。英国・イングランドでは、診療ガイドライン作成・普及に責任をもつNational Institute for Health and Clinical Excellence(NICE)の内部に“Patient Involvement Unit”という専門部局があり、診療ガイドライン作成に参加する患者代表をコーディネートするとともに、参加者に対して事前にトレーニングセッションを提供している<sup>19)</sup>。診療ガイドラインが、限られた領域の専門医だけでつくられるのではなく、社会的な役割を果たしていくためには、適切な“stakeholders”が初期の段階からの関与していることが望ましい。海外の先進的な事例を学び、日本における今後の方向性を探る手がかりを得ることは大きな意味がある。

国内では厚生労働科学研究による喘息診療ガイドライン作成班(主任研究者：宮本昭正)に、患者支援団体の一つである「アラジーポット」が参加し、患者・支援者向け情報作成に大きな役割を担

い<sup>20,21)</sup>、一般紙でも取り上げられた<sup>22)</sup>。他の領域でもこのような取り組みに関する検討がすすむことが望まれる。

稲葉は法律家の立場から、医療者と患者がともにガイドラインの限界と役割を理解し、医療者は責任と倫理を踏まえて患者の陥りやすい問題を把握し、診療ガイドラインを用いてインフォームド・コンセントを行い、対話の中で治療方針とともに決めていく調和的な医療モデルを提案している<sup>23)</sup>。患者の視点も取り入れてつくられた診療ガイドラインは、それを活用すればインフォームド・コンセントをめぐるトラブルを回避するだけでなく、インフォームド・コンセント自体を充実させ、その結果として医療の質・安全性、そして患者満足度を高めるための中核的な手段となりうるであろう。EBMの手法による診療ガイドラインが、臨床現場にとどまらず、社会的にも適切な認知と支援を受けるためには、EBMを基盤として、そこに臨床倫理や法律的な視点からの検討を加えていくことが強く望まれているといえよう。

## VI. ランダム化比較試験と観察研究の関係

EBMの普及とともに、広く知られるようになった概念の一つが「レベル・オブ・エビデンス(エビデンス・レベル)」である。個々の研究としてもっともレベルの高い、すなわち妥当性の高いエビデンスを提供するのはランダム化比較試験(RCT)であり、その次が非ランダム化比較試験、続いて観察的な疫学研究であるコホート試験やケースコントロール研究とされている。その下に位置する症例集積は多くの臨床研究報告でみられるが、対照群をもたないため、そこで観察された知見の一般化に大きな限界がある。さらに、高いレベルのエビデンスに基づかない権威者の個人的見解(エキスパートオピニオン)は、もっとも低いレベルのエビデンスとして位置づけられている。RCTの上に、個々のエビデンスを集約した2次情報であるシステマティック・レビューやメタアナリシス(データ統合型研究と呼ばれることもある)が置かれる場合もある。このように研究方法、

すなわち研究デザインによって、得られるエビデンスの科学的妥当性に上下があるという考え方は、疫学・臨床疫学においては広く認識されていたが、それらの方法論に馴染みの薄かった多くの臨床医には大きな衝撃であったことが推測される。

エビデンス・レベルに関する第一の誤解は、AHCPRによる治療の有効性に関するエビデンスのレベルが、国内のガイドラインムーブメントの初期に広く紹介されたため、治療の有効性評価以外の臨床的課題についても、それと同じエビデンス・レベルの考え方が誤って用いられてしまっている点である。Oxford EBM Centreでは、対象課題を“therapy/prevention, aetiology/harm”, “prognosis”, “diagnosis”, “differential diagnosis/symptom prevalence study”, “economic and decision analyses”の5カテゴリーに分け、それぞれの課題の特性にあった適切なエビデンスを考慮してレベルを付与している。その結果、“therapy/prevention, aetiology/harm”ではRCTのシステマティック・レビューが最上位となるが、“prognosis”, “diagnosis”, “differential diagnosis/symptom prevalence study”ではコホート研究などの観察研究のシステマティック・レビューが最上位とされている。この点については国内のガイドライン作成者も注意を払う必要がある。

エビデンス・レベルの概念の明確化を通して、研究デザインの重要性が認知されるようになったのは好ましいことではあったが、「RCTでなければならない」あるいは「RCTでありさえすればよい」というような新たな誤解が生まれたのも確かであろう。RCTは一般的になんらかの介入の有効性を評価する目的で用いられる研究デザインであるが、少なくともこれまで国内では、治験に代表されるように薬物療法の評価という文脈でもっとも受け入れられてきた。それに対して、たとえば外科領域では、2種類の方法を施術者と患者の意図を考慮せずにランダムに割り付ける介入試験を行うこと自体が実施困難な場面が多い。そこから、「外科ではEBMを行えない」という話

もときどき聞かれる。しかしEBMの実践のためには、それぞれの領域の特性を考慮して、そこで得られる最良のエビデンス、すなわち“best available evidence”を利用すればよいのであり、RCTによるエビデンスがなければできないというものではない。一般的にRCTは科学的妥当性の高いエビデンスを提示する研究デザインであるとされているが、一方で、ランダム配置に伴う倫理的な制約、厳密な組み入れ・除外基準によって選ばれた対象者での(限定的な)知見であるという一般化可能性の問題、次々と開発される新しい手技・手法の後手に回りかねない適時性の問題(ある手法の有効性についてRCTで結論が得られた時点で、その手法は陳腐化しており、次の新しい手法が現場で利用され始めている)などがある。また、費用対効果を考えれば、すべての臨床的課題をRCTで検証することは現実的ではない。一方で、エビデンス・レベルとしてはRCTに比べて低いとされている観察研究は、上記のRCTの欠点を補うものとして臨床現場で有用となる場合もある。1970～1980年代にかけて国内外で、観察研究がRCTに比べて介入の有効性を過大評価することを指摘する報告が続き<sup>24, 25)</sup>、観察研究は各種のバイアスを制御しきれないため、有効性評価の目的で用いるのは望ましくないという考え方が一般的に受け入れられた。臨床疫学、EBMの台頭とともに、その認識はさらに広く定着していったといえる。しかし、Bensonら<sup>26)</sup>、Concato<sup>27)</sup>らは近年報告された観察研究とRCTの結果をメタアナリシスの統合値として比較し、有効性の評価に大きな違いは見出されないことを報告している。観察研究(とくにコホート研究)とRCTの関係について改めて問題を提起したといえるが、それに対する反論も多い<sup>28, 29)</sup>。エビデンスをつくる立場としては、RCT、観察研究、いずれの方法を選ぶことがよいか、それぞれの臨床的課題の特性、意思決定に必要とされるエビデンスのレベル、それぞれの研究の実行可能性など、総合的に勘案して判断を行う必要がある。

## VII. 診療ガイドラインと他の2次資料との相違

海外では診療ガイドラインで推奨を示すことに異議を唱えるグループも少なくない。その代表が英国のBMJ(British Medical Journal)グループから出されている“clinical evidence”であろう。その主張は、“We supply evidence, you make the decisions. The experience of the clinical practice guideline movement has shown that it is nearly impossible to make recommendations that are appropriate in every situation.”であり、推奨を核心とする診療ガイドラインの問題点を指摘し、傾聴に値する。“Clinical evidence”のような「診療エビデンス集」は判断を(診療ガイドライン以上に)個人に委ねているが、臨床医の立場からは得意な領域では十分活用できても、経験の限られているケース、専門外の課題では、客観的なエビデンスに加えて「結局、一般的にはどうすることがすすめられるのか」という専門家の推奨が明示されたガイドラインが好まれる可能性も十分であろう。そういった意味で、診療エビデンス集と診療ガイドラインは相補的に用いられるのがよいと思われる。

“Clinical evidence”は、EBMの実践においてはいわゆる2次情報(資料)であるという点で、メタアナリシスやシステムティック・レビュー、そして診療ガイドラインと同様である。他の2次情報と対比させて、わが国における診療ガイドラインのもっている特性として次のことが考えられる。

1) 作成主体がメタアナリシスやシステムティック・レビューのように個々の研究者ではなく、学会または厚生労働科学研究班など、社会的信頼と責任など一定の「権威」をもつ組織である。

2) 一般的に作成に関与する人員が、個々のメタアナリシスやシステムティック・レビューよりも診療ガイドラインでより多く、幅広いバックグラウンドの専門家の関与が求められている。

3) 「推奨度の決定」において関係者の合意形成

過程が加わる。

4) 専門家集団内部において、個々のメタアナリシスやシステムティック・レビューとは異なり、規範としての影響力をもちうる。

5) 専門家集団内部にとどまらず、一般社会も含む専門家外部への影響力ももちうる(受診行動への影響、裁判での参照など)。

6) 専門家集団内部で一定の合意をみた見解であれば、社会的発言・政策提言という外部に向けた“professional advocacy”の足場となる。

7) 患者の立場でより信頼できる情報は、個々の(研究成果や)メタアナリシスやシステムティック・レビューではなく、社会的信頼のある主体が作成した診療ガイドラインと思われる。

## VIII. EBMとエビデンスをめぐる誤解

EBMの認知の高まりとともに、頻繁に耳にするようになったことの一つが、「EBM」と「エビデンス」の混同である<sup>30)</sup>。EBMは、①医師の専門性や経験、熟練、②患者の価値観、③エビデンスの3要素をバランスよく統合し、よりよい患者ケアのための意志決定を行うものである<sup>31)</sup>。このように、エビデンス自体はEBMの要素のうちの一つにすぎない。しかし、現在の日本では、エビデンス・レベルが高いとされるRCT、とくに大規模臨床試験の知見が得られれば「EBM」が確立し、臨床現場の意志決定までもが決まってしまうという短絡的な解釈がなされているケースがみられる。近年、数多く出版された総説・解説書でも、著名な臨床試験の解説をもってEBMとしている場合が少なくない。改めていいうまでもなく、臨床試験の知見はエビデンスの一つであり、EBMそのものではない。EBMとの混同とともにみられる「エビデンス」の過大視は、これまでの臨床的意思決定が、担当医師の経験・熟練に負う比重が大きかったことへの反動かもしれない。Haynesらは2002年のBMJ誌で“Evidence does not make decisions, people do”と強調している<sup>32)</sup>。診療ガイドラインはたとえどれほど良質につくられていたとしても、基本的には既存エビデンスの集合体であり、意思決定に影響する要因の

一つにすぎない。EBMの主唱者たちによるシンブルかつきわめて正当なメッセージを通して、国内でもこの原点を再確認することが望まれる。

## IX. 診療ガイドラインと裁判

診療ガイドラインが医療関係者のみならず、社会一般にその存在が知られていくとともに、一部の医療関係者からは診療ガイドラインが訴訟・裁判に用いられることの懸念が聞かれる。Farquharらは1990～2000年に発表された英文論文から基準に合った30篇のシステムティック・レビューにより、臨床医の診療ガイドラインに対する認識・態度を定量的に検討し、臨床医の41%が「医療訴訟を増加させる」と感じていることを報告している<sup>33)</sup>。これに対してHurwitz<sup>34)</sup>は「診療ガイドラインは現在容認されている医療慣行に則った基準を示す証拠として、expert witnessによって法廷に提示されることはあっても、専門家の証言にとってかわるものとして法廷に提示されることがあってはならない」と述べている。さらに法廷側も診療ガイドラインの内容を法律上の絶対的な基準として採択することは考えにくいとし、診療ガイドラインが存在するというだけで、いかなる状況でもそれに準拠することが妥当で、しないことが医療過誤であるということにはならないことを強調している。エビデンスとの関連がいかにも明確な診療ガイドラインであっても、個々の臨床のケースでの判断は慎重な解釈と医療者の経験に基づいた専門的判断を踏まえたうえでの適用が求められている。前出のSIGNでは、作成されたすべての診療ガイドラインに「当ガイドラインを医療基準とみなしたり、医療基準(standard)として活用したりしてはならない」ことを明記している。一方、診療ガイドライン作成者が、そのガイドラインを用いたことで個々のケースで問題が生じた場合、そのことで提訴される可能性については、法的観点からしてまず問題ないであろうとしている。通常、診療ガイドラインの作成者が個々の患者に対して配慮義務を負うことはないと考えられるからである。一方、臨床医の診療行為が診療ガイドラインから著しく逸脱している



**Minds**  
Medical Information Network Distribution Service

医療情報サービス  
厚生労働科学研究費補助金により試験公開中

Mindsについて  
Mindsの使い方  
情報提供者について

「診療ガイドライン(医療提供者向け一般向け)」をご覧になるにはこの下のボタンをクリック!!

Mindsユーザー  
ゲスト

Mindsユーザーになるには

Mindsは 無 料 でご利用になれます。登録するとメリットがあります。

○ 診療ガイドライン  
診療ガイドライン・各種医療情報はこちら

□ Mindsをお使いになる方は必ずお読みください  
診療ガイドラインをご利用になる場合について  
一般の方がご利用になる場合について

お知らせ  
『クモ膜下出血』の“Mindsアブストラクト”を公開しました  
2005.11.05 開催『第4回 EBM研究フォーラム』のお知らせ  
「Mindsユーザー」「ゲスト」ボタンをクリックしても反応しない場合はお読みください

利用条件  
Mindsをご利用になる際には、下記の利用条件についてご確認ください。

図2. Mindsのホームページ [(財)日本医療機能評価機構医療情報サービス事業]  
<<http://minds.jcqhc.or.jp/index.aspx>>

場合は、その意思決定が行われた時点でその事実と理由を患者の診療記録に十分記載する必要があることを指摘している。

診療ガイドラインへの医療専門家以外からのアクセス性が向上するとともに、診療ガイドラインの一部を取り出して拡大解釈し、訴訟のさいの根拠として用いる事例が短期的には増加する懸念はある。しかし、実際の法的判断にさいして診療ガイドラインの平均値的な記述が当該事例にそのまま適用されるものではなく、診療ガイドラインの遵守の有無で法的責任を問われるか否かが決定するものではない。医療者の立場からは、診療ガイドラインが司法の場でそのように誤用されないように、責任のある社会的発言を行っていくべきであろう。今後は「説明義務」や「期待権」の視点も考慮に入れて、診療ガイドラインの社会的普及に伴う患者との関係性の変化と、コミュニケー

ションを中心とする医療者の向き合う新たな課題について整理をすすめていく必要があると思われる。

### おわりに

1999年以降、国内ではEBMによる診療ガイドライン作成がさまざまな学会を中心に行われ、その成果が発表された。その作業過程で明らかにされた多くの課題の中で、もっとも大きなものの一つは「日本人を対象としたレベルの高いエビデンスがきわめて乏しい」事実の確認であった。基礎的な医学研究が偏重され、臨床現場の意思決定を支援する臨床研究や疫学研究が立ち遅れていた日本の研究風土の特性<sup>35,36)</sup>が、期せずして明らかにされたといえる。近年、インパクトファクターに象徴される研究業績の得点主義は、国内において臨床医を含む医学研究者の志向に少なくない影響

を与えている。インパクトファクターによる学術誌やその研究領域のランク付けと、診療ガイドラインに引用される文献(臨床の問題解決に貢献する研究と考えられる)を多く掲載している学術誌の評価は一致していない<sup>37)</sup>。EBMと診療ガイドラインの認知が高まるとともに、実際の患者の役に立つエビデンスをつくるという問題意識と適切な研究方法論が普及することで、真に意味ある臨床研究が活性化されていくことが期待される。

診療ガイドラインは医療の質を向上させるための特効薬ではない。使い方を誤れば瞬時に両刃の剣に転じる。その意義、役割、そして限界を医療者と社会一般の人々がどのように共有していくべきか、開かれたディスカッションを積み重ねていく必要があろう。診療ガイドラインに関する世界的な動きとして2002年にはWHOを含む24カ国46組織よりなる“Guideline International Net (GIN)”が発足し、2004年には推奨度の決定に関するGRADEの提案などが発表された<sup>38)</sup>。いずれも今後の診療ガイドライン問題を考えるうえで大きな意義をもつ取り組みである。ガイドラインへのアクセスビリティを高めるため、米国では国立ガイドライン・クリアリングハウス、英国では健康情報電子ライブラリーがインターネットを通じて診療ガイドラインを広く一般の人々へも公開している。国内でも現在、(財)日本医療機能評価機構の「医療情報サービス事業」、通称“Minds”(Medical Information Network Distribution Service)の整備がすすめられている(図2)<sup>39,40)</sup>。

医療に対する関心の増大、社会における医療情報のインフラストラクチャーの充実とともに、医療における患者と医師の関係の変化は加速されるであろう。EBM、そして診療ガイドラインを通じて、医療は社会における、そして社会に対するそのアカウンタビリティを問われているといえよう。

◆ ◆ ◆ 文 献 ◆ ◆ ◆

1) Field MJ, Lohr KN(ed): Guidelines for Clinical Practice ; from Development to Use, National Academy Press, Washington DC,

1992  
 2) 厚生省健康政策局研究開発振興課医療技術情報推進室(監修): わかりやすいEBM講座. 平成10年度医療技術評価推進検討会報告書, 厚生科学研究所, 東京, p8-28, 2000  
 3) 中山健夫: EBMを用いた診療ガイドライン—作成・活用ガイド, 金原出版, 東京, p78-80, 2004  
 4) 福井次矢, 丹後俊郎: 診療ガイドラインの作成の手順. EBMジャーナル4: 28-37, 2003  
 5) 棚橋紀夫: EBMに基づく治療ガイドラインの作成と応用の問題点. 第44回日本神経学会総会プログラム・抄録集, 日本神経学会, 東京, p61, 2003  
 6) Last JM: A Dictionary of Epidemiology, 4th Ed, Oxford University Press, Oxford, 2000  
 7) Nakayama T, Budgell B, Tsutani K: Confusion about the concept of clinical practice guidelines in Japan ; on the way to a social consensus. Int J Qual Health Care 15: 359-360, 2003  
 8) Eddy DM: Clinical decision making ; from theory to practice. Designing a practice policy. Standards, guidelines, and options. JAMA 263: 3077, 3081, 3084, 1990  
 9) 津谷喜一郎, 長澤道行: 医師と診療ガイドライン—“professional autonomy”の視点から. 日医師会誌129: 1793-1803, 2003  
 10) Grimshaw JM, Russell IT: Effect of clinical guidelines on medical practice ; a systematic review of rigorous evaluations. Lancet 342: 1317-1322, 1993  
 11) Koyama H, Fukui T: A review of research on the clinical effectiveness of therapeutic practice guidelines, 1991~2000 ; the need for standardization of reporting style. General Med 1: 1-7, 2002  
 12) Appraisal of Guidelines, Research, and Evaluation in Europe(AGREE) Collaborative Group: Guideline development in Europe ; an international comparison. Int J Technol Assess Health Care 16: 1039-1049, 2000  
 13) 長谷川友紀: 診療ガイドラインを取り巻く状況—AGREE collaborationの動向. EBMジャーナル4: 294-297, 2003  
 14) Shiffman RN, Shekelle P, Overhage JM et al: Standardized reporting of clinical practice guidelines ; a proposal from the Conference

- on Guideline Standardization. *Ann Intern Med* 139 : 493-498, 2003
- 15) 中山健夫 : EBMを用いたガイドラインの作成と応用—EBMの考え方と問題点を踏まえて. *Prog Med* 23 : 3143-3151, 2003
  - 16) Lenzer J : Alteplase for stroke ; money and optimistic claims buttress the “brain attack” campaign. *BMJ* 324 : 723-729, 2002
  - 17) Choudhry NK, Stelfox HT, Detsky AS : Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 287 : 612-617, 2002
  - 18) Nomura H, Nakayama T : The Japanese healthcare system ; the issue is to solve the “tragedy of the commons” without making another. *BMJ* 24 : 648-649, 2005
  - 19) 鈴木博道 : NICEガイドライン開発への患者・介護者の参画. *あいみっく* 25 : 10-14, 2004
  - 20) 宮本昭正 (監) : EBMに基づいた患者と医療スタッフのパートナーシップのための喘息診療ガイドライン2004—小児編, 協和企画, 東京, 2004
  - 21) 宮本昭正 (監) : EBMに基づいた患者と医療スタッフのパートナーシップのための喘息診療ガイドライン2004—成人編, 協和企画, 東京, 2004
  - 22) ぜんそく診療—患者参加で指針作り, 分かりやすく図解. 朝日新聞, 2004年6月13日朝刊
  - 23) 稲葉一人 : インフォームド・コンセントを充実させるためのガイドライン. 厚生労働科学EBMを指向した「診療ガイドライン」と医学データベースに利用される「構造化抄録」作成の方法論の開発とそれらの受容性に関する研究 (主任研究者・中山健夫)2002年度報告書, p50-58, 2003
  - 24) 砂原茂一 : 抗炎症薬に関する有効報告—RCTとhistorical controlled studyの比較. *日内会誌* 59 : 588, 1970
  - 25) Sacks H, Chalmers TC, Smith H Jr : Randomized versus historical controls for clinical trials. *Am J Med* 72 : 233-340, 1982
  - 26) Benson K, Hartz AJ : A comparison of observational studies and randomized, controlled trials. *N Engl J Med* 342 : 1878-1886, 2000
  - 27) Concato J, Shah N, Horwitz RI : Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 342 : 1887-1892, 2000
  - 28) Ioannidis JP, Haidich AB, Pappa M et al : Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 286 : 821-830, 2001
  - 29) Pocock SJ, Elbourne DR : Randomized trials or observational tribulations? *N Engl J Med* 342 : 1907-1909, 2000
  - 30) 中山健夫 : EBMの手法を用いたガイドラインの作成と普及に向けて. *Ther Res* 26 : 305-311, 2005
  - 31) Haynes RB, Devereaux PJ, Guyatt GH : Clinical expertise in the era of evidence-based medicine and patient choice. *ACP J Club* 136 : A11-14, 2002
  - 32) Haynes RB, Devereaux PJ, Guyatt GH : Physicians' and patients' choices in evidence based practice. *BMJ* 324 : 1350, 2002
  - 33) Farquhar CM, Kofa EW, Slutsky JR : Clinicians' attitudes to clinical practice guidelines. *Med J Aust* 177 : 502-506, 2003
  - 34) Hurwitz B : Legal and political considerations of clinical practice guidelines. *BMJ* 318 : 661-664, 1999
  - 35) Fukui T, Rahman M : Contribution of research in basic and clinical sciences in Japan. *Intern Med* 41 : 626-628, 2002
  - 36) Nakayama T, Yamazaki S : Percentages of reports of clinical trials, written in seven non-English languages, that have structured abstracts. *General Med* 4 : 7-10, 2003
  - 37) Nakayama T, Fukui T, Fukuhara S et al : Comparison between impact factors and citations in evidence-based practice guidelines. *JAMA* 290 : 755-756, 2003
  - 38) Atkins D, Best D, Briss PA et al : GRADE Working Group ; grading quality of evidence and strength of recommendations. *BMJ* 328 : 1490, 2004
  - 39) Minds入門—インターネットによる診療ガイドライン活用方法. *医事新報* 4184 : 1-15, 2004
  - 40) Satoh T, Nakayama T, Sato Y et al : Physicians' awareness regarding evidence-based medicine, practice guidelines and clinical information resources in Japan ; Needs assessment prior to the initiation of “Medical Information Network Distribution Service (Minds)”. *General Med* 5 : 13-20, 2004

# 研究計画書

## 関節リウマチ診療ガイドラインの 診療パターンに対する影響に関する研究 日本リウマチ財団登録医を対象とした質問票調査

主任研究者： 京都大学医学研究科 健康情報学分野 中山 健夫  
共同研究者： 独立行政法人国立病院機構相模原病院 越智 隆弘  
共同研究者： 京都大学医学研究科 医療疫学分野 福原 俊一  
共同研究者： 京都大学医学研究科 医療疫学分野 東 尚弘

2006年1月25日