

4. 診療ガイドラインがカバーする範囲

患者の 60-95%とされる (Eddy)。95%以上の患者に適応されるのは「スタンダード」、50%ほどの患者には「オプション(選択肢)」が適用されることもある (7)。

5. 診療ガイドラインの評価法

AGREE 共同計画 (“Appraisal of Guidelines Research and Evaluation”) が開発した 6 領域 23 項目と総合評価、合計 24 項目から成る評価手法がある (8,9)。無批判に診療ガイドラインを利用するのではなく、その限界を理解して適切に利用することが望まれる。

* 下記の各項目を 4 段階で評価する。利用法については下記を参照。

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

(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。

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の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。

(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。

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。

(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。

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。

(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。

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。

(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。

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。

(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。

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。

(7)総合評価: 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?

6. 利害の衝突(conflict of interest)

相反利益(competing interest)とも言う。医学雑誌編集者委員会(International Committee of Medical Journal Editors: ICMJE)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(例えば雇用、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、通常最も重大な「利害の衝突」としてみなされる。診療ガイドラインやその関連資料の利用に際しては、作成者・執筆者と特定の製薬企業との関係について注意する必要がある(10,11)。

7. 患者・介護者とのコミュニケーションの基点としての診療ガイドライン

稲葉は法律家の立場から、医療者と患者・介護者が共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・コンセントを行ない、対話の中で治療方針を共に決めていく(shared decision making)調和的な医療モデルを提案している。診療ガイドラインが患者の視点を取り入れて作られ、個々の臨床場面で適切に用いられれば、インフォームド・コンセントを巡るトラブルの回避だけではなく、インフォームド・コンセント自体を充実させ、医療の質・安全性、そして患者満足度を高めるための重要な手段となり得る(12)。

8. エビデンスが明確でない場合の意思決定の基準

予防医学に関するカナダ・タスクフォースは、以下の基準を提案している。

- 意思決定における患者自身の参加を促す
- 害を最小化する
- 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
- 不要な「ラベリング(病気であるとレッテルを貼ること)」を避ける
- 益の不確かな高価な手技を避ける
- 疾病負担が大きい状況に焦点を当てる
- ハイリスクグループの特別のニーズに配慮する

9. 診療ガイドラインと訴訟

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、個々の臨床場面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上での適用が不可欠である。診療ガイドラインに書いてある通りのことをすれば良い、書いてないことをすることは許されず訴訟の対象になるというものではない。しかし一方では、診療行為が診療ガイドラインから著しく異なっている場合は、その事実と理由を診療記録に

記載する必要性が高まっていくと予想される 13)。

10. 診療ガイドラインの情報源

各種の診療ガイドラインは出版物として利用できる他、インターネット上でも公開が進んでいる 14)。

*インターネット上から利用できる診療ガイドライン

- 各学会ホームページ
- 財団法人日本医療機能評価機構・医療情報サービス事業による”Minds” (Medical Information Network Distribution Service)。 <http://minds.jcqh.or.jp/to/index.aspx>
- 東邦大学医学メディアセンター <http://www.mnc.toho-u.ac.jp/mmc/guideline/>
- 医学中央雑誌 Web <http://www.jamas.gr.jp/>

(文献)

1. Institute of Medicine. Guidelines for clinical practice: from development to use. Washington DC, National Academy Press, 1992.
2. 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, 2004
3. 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 2003;4:28-37.
4. 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. 日本医師会雑誌 129:1793-803, 2003
5. J. Last. A Dictionary of Epidemiology (4th edition), Oxford University Press, Oxford, 2000
6. Haynes RB, Devereaux PJ, Guyatt GH. Physicians' and patients' choices in evidence based practice. BMJ 2002;324:1350.
7. Eddy DM. Clinical decision making: from theory to practice. Designing a practice policy. Standards, guidelines, and options. JAMA 1990; 263: 3077, 3081, 3084.
8. Appraisal of Guidelines, Research, and Evaluation in Europe (AGREE) Collaborative Group. Guideline development in Europe: An international comparison. Int J Technol Assess Health Care. 2000;16:1039-49.
9. 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 2003;4:294-7.
10. Lenzer J. Alteplase for stroke: money and optimistic claims buttress the “brain attack” campaign. BMJ 2002;324:723-9.
11. Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. JAMA 2002;287:612-7.
12. 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
13. Hurwitz B. Legal and political considerations of clinical practice guidelines. BMJ 318:661-664, 1999.
14. 特集 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 2004;4184:1-15.

【研究論文：依頼原稿】

エビデンスに基づく医療（EBM）の系譜と方向性： 保健医療評価に果たすコクラン共同計画の役割と未来

正木 朋也

国際医療福祉大学大学院保健医療学専攻
masakit-tky@umin.ac.jp

津谷 喜一郎

東京大学大学院薬学系研究科
tsutani-tky@umin.ac.jp

要 約

国内外の公共投資について科学的なエビデンス¹を明らかにすべきとの社会要請に対し、日本における経験は保健医療分野を除いて皆無と言わざるを得ない²。しかし、イギリス、アメリカにおいては、公共政策判断の場においてもエビデンスに基づく評価が行われ、着実にエビデンスが蓄積されつつある。キャンベル共同計画、アメリカ教育省のWhat Works Clearinghouse、ODA分野の貧困アクションラボの事例などがそれにあたる。このような背景を踏まえ、本稿では、まず先行する「エビデンスに基づく医療」(Evidence-based Medicine: EBM)の系譜とコクラン共同計画について概観する。ついで、社会基盤としての質の高い情報の蓄積の重要性と社会認識の問題を踏まえ、EBMの今後の方向性について考察した。

キーワード

EBM、エビデンス、ランダム化比較試験、システマティック・レビュー、社会認識

はじめに

国内外の公共投資が効果的に実施されているかを考えるとき、プロジェクトやプログラムレベルにおける効果測定や判定も重要であるが、社会的コンテキストを踏まえたよりマクロな政策の評価もまた重要である。本稿では、このような視点から「エビデンスに基づく医療」(Evidence-based Medicine: EBM)と政策評価の系譜についての考察を試みる。まず、EBMという概念が誕生した歴史背景を概観し、その社会的要請がいかに達成されたか(されつつあるか)をみる。次いで、情報インフラの役割を担っているコクラン共同計画(The Cochrane Collaboration: CC)³について、シス

テマティック・レビュー(Systematic Review: SR)とその結果をまとめてデータベースとしたコクラン・ライブラリについて概説する。さらに、情報の質の確保について要点をレビューし、研究デザインと批判的吟味、研究結果の要約を効率よく伝えるための構造化抄録の重要性などについて述べる。

因みに、EBMは広義には、Evidence-based Healthcare、Evidence-based Public Health、Evidence-based Policy-making他、保健医療のみならず健康に関わるあらゆる領域を含む。一方、狭義には、文字通り、まさに一人の患者を対象とした医療現場での治療に関わる狭義のプラクティスを示す本来の意味としてのEBMがあり、その関

連書籍も多い。本稿では上述の視点から、後者の詳細については殆ど触れず、公共投資の評価とマクロレベルの政策に関わる視点からEBMを紹介させて頂くことをご承知頂きたい。本来、前者は広義のEBMとしてEvidence-based Healthcare (EBH) とでも表記し区別すべきものであるが、現実には、後者(狭義のEBM)をも含めた(合わさった)概念としてEBMという用語が普及・定着しつつある。EBMという用語を初めて目にされた方は、この状況に留意のうえで本稿をお読み頂ければ幸いである。

このようにEBMには、広義と狭義の意味があるが、その根底には対象となる患者の好み(patient preference)、価値システム(value systems)、および現場の状況などをも踏まえた共通性がある。時に「ランダム化比較試験(randomized controlled trials: RCT)に基づく医療」(RCT-based Medicine)との批判を受けることもあるが、その認識は誤りである。EBMの基本は、研究によるエビデンスの産出にあり、そのエビデンスにはグレードがあることを明白にしたうえで、それを場の状況に応じて適切に利用することである。

実績と経験にも支えられたEBMの考え方は極めて有益であるので、他の分野・領域においても、その現状に則した評価活動に活かすことが可能である。

本稿ではさらに、この基本的な考え方を支える科学の公衆的理解(public understanding of science: PUS)の重要性にも触れ、適切なナレッジ・マネジメントの仕組みが、基本的な社会要請(事実を正確に公表した上で公明正大な判断を行う)に対するひとつの有力な回答となり得ていることを述べ、社会認識(social epistemology)に及ぼす影響と未来展望についても言及した。

1. 政策評価とEBM

Evidence-based Medicine (EBM) という用語は、カナダのマクマスター大学のGuyattの論文(Guyatt 1991)に用いられたのが始まりとされる。EBMという用語はその後、特にSackettらの“Evidence-based Medicine: how to practice and teach EBM”

(Sackett 1996)発行後、さらに世界的に用いられるようになった。

日本では、「科学的根拠に基づいた医療」などと訳されることもあった。上述の広義・狭義の違いに加え、「科学的根拠」という一般的用語として理解されることによる議論が生じ易いため、昨今著者らは「エビデンス」というカナ標記を用いることを推奨している。

ここでは、政策評価の変遷を振り返りながら、EBMがどのような時代背景に生まれ育ったかについて歴史背景をもとに考察する。

(1) イギリスの研究開発戦略とエビデンスに基づく保健医療 (EBH)

イギリスでは1980年代の行政改革の動向を背景に、医学研究に対する深刻な論議が起こった(Smith 1988)。当時の論点は、①研究活動の低下、②基礎研究と応用研究のバランスの崩れ、③研究結果の未利用、であった。上院が委員会を設置し問題を検討した結果、健康サービス研究(health service research)と公衆衛生研究(public health research)を促進すべきという結論に至った。当時としては、大変驚くべき結論であったが、政府はこれを歓迎した。

こうした研究の方向性の転換は、英国保健省(National Health Service: NHS)のサービス提供と国家的な研究投資の費用-効率に大きな影響をもたらした。1990年に政府はNHSの中核に研究開発部門を設置して1991年には保健医療の研究開発戦略を打ち出し、①国民の健康・福祉の改善、②科学的な情報に基づく医療政策と医療行為の促進、③効果的で効率的な資源の利用、のための行政改革を実施した。これらの改革は、それまでの意思決定者の意見に基づく政策決定から、エビデンスに基づいた政策決定へとシフトするきっかけ作りになったとともに、政策実施により成果を挙げたかどうか、納税者である国民に対しその説明責任を果たす試みでもあった(ミユア・グレイ1999⁵⁾。この改革の目的は、エビデンスに基づく保健医療(Evidence-based Healthcare: EBH⁶⁾そのものといえる。

このような背景の中、1992年10月に研究開発部

門の情報ネットワークと協力する形でThe Cochrane Centreがオックスフォードで活動を開始し、翌1993年3月にThe UK Cochrane Centreと名称変更の後、10月の第1回コクラン・コロキウム (Oxford) により正式にCCとしての活動を開始した。また同年にはレビュー作成と共にイギリスの医療関係者をはじめとして、情報を必要とする人々にレビューとその結果の意味を普及させる目的で、政府直轄のNHSレビュー普及センター (NHS, the Centre for Reviews and Dissemination: NHS CRD) がヨーク大学内に設置された。

イギリス政府の行政改革や世論としての研究開発思考の追い風もあり、計画的かつ継続的に質の高いエビデンスを維持・更新する基盤が確立した。アメリカの研究者や行政評価の立役者らとも相互に呼応して、その後、十数年にわたる情報と経験の蓄積を経て、ようやく今日のゆるぎない保健医療技術評価基盤を形成し得たものと理解できる。広義の概念を擁するEBHの発展は、それぞれ保健制度の異なるイギリスとアメリカにおいてほぼ同時に進展し、情報技術のめざましい発展の恩恵も受けながら、次第に世界中で受け入れられ、多くのボランティアらの協力を得て現在もなお拡大展開しつつ今日に至っている。

(2) アメリカを中心とした政策評価の変遷とEBM

アメリカの政策評価の歴史を振り返ると1930代の教育と公衆衛生の分野にその原型をみることができる (Rossi et al 1999)。1965年合衆国連邦政府は、ジョンソン大統領が打ち出した2つの政策的スローガン (①貧困との戦い、②偉大なる社会の建設) 達成のため全ての省庁に予算配分システムPPBS (Planning Programming Budgeting System) を導入し、事業目的、成果評価、代替案の作成、効率性評価という一連の手続きの下に予算配分を行った。このPPBSの影響で、これまであまり脚光を浴びることの無かった政策評価についても、その評価手法や方法論に関する研究と実績が飛躍的に増えることとなった。なお、PPBSは1960年代後半に合衆国連邦政府で導入が開始され、地方政府も含めて一応の定着を見たプログラム評価

(今日的な意味での政策評価) の起源ともいうべきものであったが、その定着は失敗したとの見解がある (窪田1998)。

その後、1980年代に入り、財政赤字の抑制を目的とした公共プログラムの絞込みと効率化のための道具として政策評価が採用されるようになり、その結果この時期には、評価手法・評価研究それ自体の質的な向上と発展がみられた。

1990年代は市場経済の波に押され、政策評価についても財務的・経済的な効率性の比重が高くなり、科学的な評価からより実用的な評価へと評価対象が広がっていった。連邦政府は1992年に行政改革の綱領ともいえるNational Performance Review (NPR) を発表し、①より少ないコストで、よりよい成果を上げる、②消費者を最優先に、③手続き重視から結果重視へ、を掲げ一般に、より結果が理解され易い評価を重視する機運が醸成された。1993年には「政府業績成果法」(Government Performance and Results Act: GPRA) も成立し、事前・事後の比較等を応用し、計量可能な指標を用いたパフォーマンス・メジャーメントによる評価が広範に利用され始めるに至った (龍、佐々木2000、田辺2003)。

このような背景の中、連邦政府は医療分野に関してもGPRAの下位に保健医療法 (Health Care Act) を制定し、NPRの具体的項目を達成するために必要な法的整備を行った。NPRでは行政施策が何を目指し、どのような手続きを経て、どのような目的を達成したかをみることにしていた。その手法は、Evidence-based Management、Result Oriented Management、Performance-based Management、Managing for Resultsなどさまざまな用語で呼ばれているが、何れも経営管理論の始祖として知られるアンリ・ファヨール (Henri Fayol, 1841-1925) の示した“Plan, Do, See”サイクルの循環の上に成り立つ。即ち、事前に定めた計画に従って、目的に合致した活動を実行し、その効果を測定し評価するという、改善に向けての基本的かつ連続的なアプローチであった。

John Rodman Paul (1893-1971) が1938年に米国臨床研究学会で提唱し (Paul 1938)、Alvan R. Feinstein (1925-2001) により始められた臨床疫学 (Feinstein 1968) におけるOutcome Researchや

Decision Analysis、Cost-Effectiveness Analysisなどは、上述のマクロレベルの社会改革の進展に呼応した医療評価におけるアプローチといえる (Sackett 1991)。その背景には、経営管理医療 (マネジド・ケア) や健康維持組織 (health maintenance organization: HMO) による医療の効率化とNPRの医療適用について、当時北米を席卷していた社会改革の手法が少なからず影響したであろうことは容易に想像できる。

重要な点はEBMそれ自体が独立先行し現在に至ったわけではなく、むしろ社会改革の一翼を担い、また、社会の要請に応える形で、他の社会改革評価手法と共鳴しつつ発生・発達し得たことである (林2000)。

2. コクラン共同計画とシステマティック・レビュー

(1) エビデンスを「つくる」「つたえる」「つかう」

コクラン共同計画 (The Cochrane Collaboration: CC) は1992年にイギリスの国民保健サービスの一環として始まり、現在世界的に展開されている医療技術評価のプロジェクトである。ランダム化比較試験 (randomized controlled trial: RCT) を中心に、世界中の臨床試験のシステマティック・レビュー (SR, 収集し、質評価を行い、統計学的に統合する) を行い、その結果を、医療関係者や医療政策決定者、さらには消費者に届け、合理的な意思決定に供することを目的とし、EBMの情報インフラストラクチャーとしての役割を担っている。

コクランというのはアーチボルト・コクラン (Archiebald Cochrane, 1909-1988) という医師で疫学者の名であり、彼は1979年に「各トピック毎に、それぞれのランダム化比較試験を、定期的にクリティカルにまとめていないことに関して、われわれ専門家は批判されるべきである。」と述べた。

この意味するところは、①すでにあるRCTから、②よいものを選びすぐり質の悪いものは省いて、③それらをまとめて、④遅滞なく、⑤必要な人に届ける、ことが大切だということである。

エビデンスは「つくる」「つたえる」「つかう」ことに意義があり、それに関わる立場も3つある (表1)。コクランの発言で最も重要な点は、エビデンスを「つくる」場と「つかう」場の間に入って「つたえる」ことが大事だと力説したことである。なお、エビデンスは「つかって」こそ意味があるのだが、実際にはこの「つたえる」ことは意外に容易でないことについては後で述べる。

表1 エビデンスに関わる3つの立場

つくる	— 臨床試験
つたえる	— コクラン共同計画
つかう	— 臨床家、行政官、消費者

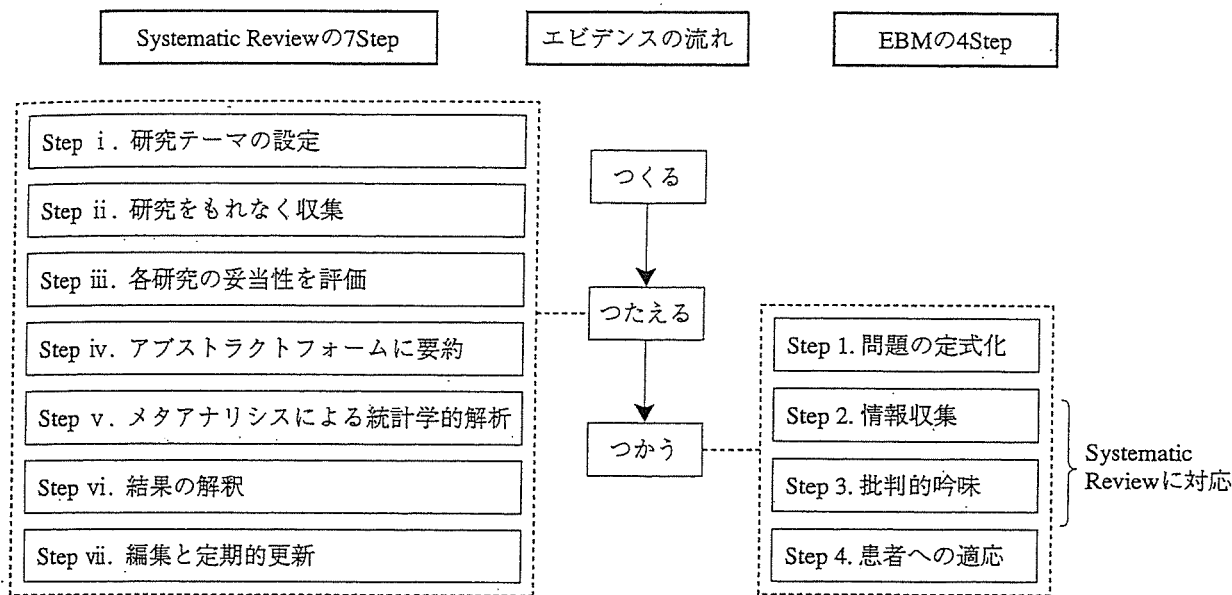
(出所) 津谷2000a

(2) システマティック・レビュー (収集し、質評価を行い、統合する)

CCはコクランの弟子のIan Chalmers というイギリスの産婦人科医が始めたものである。彼はコクラン共同計画を始める前から、10年以上にわたり周産期領域のシステマティック・レビュー (SR) を行っていた。CCの「つたえる」ことの前半部分 (情報の生産) の中心的作業がSRである。後半部分の作業は利用者へのサービス提供であり、そのプロダクトが後に述べるコクラン・ライブラリである。

SRの具体的手順は図1のステップによるが、手順の詳細とCCが提供するシステマティック・レビュー・ハンドブックなどの概要については既報 (津谷1997, 2000a, 2000b) を参照されたい。ここで強調しておきたいことは、部品としてのRCTを実施すること (あるいはエビデンスをつくること) 自体はもちろん重要であるが、既存の蓄積された研究成果 (領域によっては必ずしもRCTがなされているとは限らない) を、批判的・系統的に吟味し、重要な結果のみを抽出・整理・上清してゆくSRのプロセスこそが最も重要であるという点である。即ち、SRはCCをはじめ、キャンベル共同計画、アメリカ教育省のWhat Works Clearinghouse および貧困アクションラボなどに共通する活動の本質ともいえる。

図1 システマティック・レビュー



(出所) <http://www.lifescience.jp/ebm/opinion/200308/#fig01> (accessed Jan 3, 2006)

表2 コクラン・ライブラリ

Database	Total Records
1) The Cochrane Database of Systematic Reviews (Cochrane Reviews)	4,113
2) Database of Abstracts of Reviews of Effects (DARE)	5,585
3) The Cochrane Central Register of Controlled Trials (CENTRAL)	463,763
4) The Cochrane Database of Methodology Reviews (Methodology Reviews)	20
5) The Cochrane Methodology Register (Methodology Register)	7,484
6) Health Technology Assessment Database (HTA)	4,733
7) NHS Economic Evaluation Database (NHS EED)	16,418
8) About The Cochrane Collaboration and the Cochrane Collaborative Review Groups (About)	90

(出所) The Cochrane Library, <http://www.thecochranelibrary.com/> (accessed Jan 3, 2006)

なお、コクラン共同計画では上記全体のプロセスそのものをシステマティック・レビューと呼び、統計学的・数学的なプロセスの部分をメタ・アナリシスと呼んでいるが、一般には全体をメタ・アナリシスと呼んでも差し支えないと思われる。昨今では、時に、質的研究のメタ・アナリシスという言葉も耳にするが、この場合は既存研究を統合する研究の総称であり、さらに広い意味でメタ・アナリシスという用語が用いられている。EBM同様、狭義と広義の意味の違いには注意が必要である。

(3) コクラン・ライブラリ

先にCCの主要な活動の「つたえる」ことには、前半部分（情報の生産）と後半部分（利用者へのサービス提供）があることを述べた。その後半部分のプロダクトがコクラン・ライブラリである。CCでは、上述のSRを経て四半期に一度コクラン・ライブラリを更新し有料でその成果を一般に還元している。CD-ROMに専用の検索ソフトをバンドルし販売するとともに、ホームページからWebによる利用も可能である。日本の医学系学部をもつ多くの大学で利用可能であり、個人などで

の購入も可能である⁸⁾。

コクラン・ライブラリは8種類のデータベースからなり、2006年1月時点でそれぞれ表2に示す件数の登録がある (The Cochrane Library 2005, Issue 4)。

それぞれのデータベースの内容は以下のとおり。

- 1) The Cochrane Database of Systematic Reviews (Cochrane Reviews)
コクラン・ライブラリの中心的データベース。各種疾患を対象とした治療や診断についてのSRを収録し、Complete reviewsとProtocols (レビュー中のもの) からなる。CCのCollaborative Review Group (CRG) が“Cochrane Reviewers' Handbook”に基づき作成。
- 2) Database of Abstracts of Reviews of Effectiveness (DARE)
学術誌等で公表されたSRを収録。イギリスのヨーク大学NHS Center for Reviews and Dissemination (CRD) において、基準を満たすレビューを選択し、構造化抄録を作成。
- 3) The Cochrane Central Register of Controlled Trials (CENTRAL)
RCT検索の中心的データベース。ランダム化比較試験や比較臨床試験 (準ランダム化) の論文、学会抄録等の書誌情報を収録。非公開の臨床試験も収録している。
- 4) The Cochrane Database of Methodology Reviews (Methodology Reviews)
SRの方法論を検討しているMethodology Review Group 作成のデータベース。
- 5) The Cochrane Methodology Register (Methodology Register)
SRに関する方法論 (作成手法・ランダム化比較試験の研究手法) 論文の書誌情報のデータベース (EBM実践の為の資料集)。
- 6) Health Technology Assessment Database (HTA)
医療技術評価の文献を収録。治療、予防、リハビリテーション、医療機器、手術手法などを対象。NHS CRDにより維持管理。

- 7) NHS Economic Evaluation Database (NHS EED)

医療の経済効果、費用分析論文の構造化抄録・書誌情報データベース。NHS CRDにより維持管理。

- 8) About the Cochrane Collaboration and the Cochrane Collaborative Review Groups (About)
コクラン共同計画そのものの情報。

表2上の8) About The Cochrane Collaboration and the Cochrane Collaborative Review Groups (About) をみると、ミッション・ステートメントとともに、各運営委員会メンバーが過去5年以内にどこから資金援助を受けたか、コンサルテーションに関わる報酬を得たか、などの情報が公開されており、独立性と中立性を自ら示そうとする姿勢も興味深い。

また、Collaborative Review Groups (CRGs) のリストもあり、終了したReviewや現在進行中のProtocolの全てを一覧することができる。

なお、DARE、HTA、NHS EED は、NHS CRDのDAREサイト⁹⁾にて無料で検索可能である。

Cochrane Reviewのabstractについても無料でweb上で公開されている。その日本語訳は日本のネットワークであるJANCOCでボランティアになされていたが、作業量の多さなどから中断した。その後、(財) 日本医療機能評価機構のMINDSで診療ガイドラインに関連したCochrane Reviewを中心に訳がなされ、2006年3月から無料公開がスタートした。

- (4) 組織体制、レビュー・グループ、関連ネットワーク

CCは独立した中立機関として活動を展開し、現在12のセンター (各支部を加えると計25) が地域毎の教育や活動支援を行うとともに、レビューの重複が生じないように調整と維持管理にあたっている。

これらのセンターを拠点に、現在50のレビュー・グループが活動しておりその殆どが独自のホームページを持ち、進行中のレビュー内容について相互に参照できるようにしている。レビュー・

グループの殆どは疾病・治療領域毎に形成されているが、中にはMethodology Review Groupのように、レビュー方法そのものについての検討グループもある。

CCではさらに、異なる研究グループとの協力体制をとるため13の分野（カウンターパート／専門グループを含む）について担当者を定め公表している。キャンベル共同計画や各領域の専門家グループらと相互に連携することにより、常に最新の動向を共有するとともに、重複したレビューの調整などを行うことが目的である。同様の相互連携や協力体制はキャンベル共同計画の教育グループとアメリカ教育省のWhat Works Clearinghouseとの間にもあり、目的と方法論を共有できるグループ同士が相互の活動を尊重しつつ連携をとることは、SRを推進する立場や過去の歴史的な背景から考えても極めて自然な流れといえる。

3. 情報の質の確保

インターネットの世界的な普及により、情報量は指数関数的に増加している。このように情報が氾濫する中で、我々はどのように情報を取捨選択し、合理的な判断を行うことができるだろうか？このような、情報の質の確保とその評価は、行政判断あるいは個人の生活に直接影響する問題であるにも拘わらず、多くの場合なおざりにされている。情報が日々刻々と変容することを考えれば、この問題は受身の姿勢では解決できないことは明らかである。ここでは、情報の質をいかに積極的に高く維持するかという観点から、エンドポイント、バイアス、試験デザイン、エビデンスのレベルについて留意事項を述べる。

(1) エンドポイント (endpoint)

臨床研究において用いるエンドポイントという用語は、評価目標、評価項目、評価基準、評価尺度、評価のものさし、といった意味であり（津谷1995）、ICH統計ガイドライン（E9）¹⁰の定訳として「主要変数（「目標」変数、主要評価項目ともいう）」がある。英語ではindicator、marker、

parameterなどと呼ばれるもので、臨床評価や薬効評価においては、時に、主要評価項目（primary endpoint）副次的評価項目（secondary endpoint (s)）などにより介入効果を測定し、効果判定のよりどころとする。要は当該研究で「何を見たいか」をエンドポイントとし、一番重要な項目を主要評価項目と定義して重要視し、それ以外のものを副次的評価項目として参考とする。

主要評価項目が疾病による生死といったように観察しづらい場合には、その代替エンドポイント（surrogate endpoint）を用いる場合もある。これは真のエンドポイント（true endpoint）に対する用語である。例えば糖尿病や高脂血症の患者を治療対象として考えるとき、本当は薬剤投与の効果による延命（生死が真のエンドポイント）を臨床研究の結果としてみたいのであるが、倫理的、経済的、および時間的制約などの実施可能性から生死を主要評価項目とすることができないため、過去の研究などから疾患の病態と相関することが明らかになっている血糖値やコレステロール値を代替エンドポイントとして臨床試験を実施し、本来知りたいエビデンスを間接的に産出する。なお、代替エンドポイントが正しく選択されていたかについては、通常、平行あるいはその後実施される疫学調査などにより、後になって明らかにされることになる。

エンドポイントは「客観性objectivity」の高いものとすべきとの考え方が強く、数値や画像で表されるものや、死亡のように「あり」「なし」で表されるものは明確にみえるので、ハード・データと呼ぶ。一方、生活の質（Quality of Life: QOL）のように本人の「主観性subjectivity」や時に「第三者の主観性」を重視したものをソフト・データと呼び区別する場合もある。

何れの場合でも、エンドポイントは客観的であることが望まれる。しかし、「主観性」はもちろんであるが「客観性」にも社会文化的な価値観が反映されるため「客観性」のあるエンドポイントを設定することは必ずしも容易ではない。特に文化的価値観に相違のある歴史的あるいは国際的な研究を行う場合には「客観性」という言葉が常に曖昧さを含んでいることには十分留意する必要がある。

なお、「信頼性reliability」と「再現性reproducibility」は結果的にほぼ同義と考えることができる。観察者や測定日時が変わった場合でも同様の結果を得るために「信頼性（再現性）」の確保が重要である。「信頼性（再現性）」が確保されていればその結果の評価は容易になる。

しかし、次に述べるような、様々な原因により発生するバイアス（bias）に妨げられて再現性が確保できないことは評価の現場ではよく経験されることである。以上を踏まえれば、科学的なエビデンスを産出するためには上述した問題を十分意識したうえで、積極的にこれらバイアスを排除しなければならないことが改めて理解できる。

(2) バイアスと研究の事前登録

あらゆる研究や評価には必ず何らかのバイアスが入り込むと考えるべきである。統計学では傾向のないランダム誤差（random error）であるバラツキ（variability）に対して、バイアスは系統誤差（systematic error）として捉えられる。

マクマスター大学のハダットの著書（ハダット2004¹¹）には36種類ものバイアスが紹介されている。1982年に報告されたもの（Owen 1982）を補足し「利用者が情報を理解する過程で生じるバイアス」として紹介している。ここで各バイアスの詳細説明は省略するが以下の括弧内の語の後にバイアスという語をつけ、語感から想定されるバイアスの内容を推察すると面白い。

「競争」「彼には借りがある」「モラル」「専門領域」「何かをしなければならない」「好みのデザイン」「有名雑誌」「有名な著者」「大規模試験」「はでなタイトル」「資格あるいは専門的なバックグラウンド」「地理的」「公表物の言語による」「怠慢」「伝統」「銀行通帳」「反抗」「テクノロジー」「経験主義」「私は疫学者である」

有名の反対には無名、大規模の逆に小規模といった表裏をなすバイアスがあることを考えれば、さらに多くのバイアスが存在することが理解できる。バイアスというと、研究デザインを立案するときに可能な限り取り除くべきものであるといっ

たことを考えがちであるが、得られた結果を解釈する際にも気づかぬうちに多くのバイアスを自ら抱え込んでいることは認識しておくべきであろう。

また、研究者やスポンサー側からみて良い（あるいは悪い）結果だけを報告するといった出版バイアスにも注意が必要である。こうした問題に敢えて立ち向かうことを表明し2002年に発行を開始した基礎医学の電子雑誌Journal of Negative Results in Biomedicine¹²は、出版バイアスが研究全体の方向性に誤った示唆を与え、ひいては社会認識に問題を生じさせることに対する警鐘を与えている。

2004年9月、国際医学雑誌編集者委員会（International Committee of Medical Journal Editors: ICMJE）は、臨床試験の事前登録がなされていることを雑誌投稿の条件とした。翌月カナダ・オタワにて、国際的な臨床試験の登録をサポートする原則を構築し、さらにその実施を推進することを目的とした会議が催された。これを受け、世界保健機関（WHO）は、2005年にInternational Clinical Trials Registry Platform（ICTRP）¹³プロジェクトを立ち上げた。これは世界に多数存在する臨床試験登録システムを、一定基準の下に束ね、これらの臨床試験登録システム内で一意となる試験識別番号を発行することにより、世界中に散在する臨床試験情報を統一的に検索できるようにすることを目指した活動である。

日本においては、2005年6月に「UMIN臨床試験登録システム」（UMIN Clinical Trial Registry: UMIN-CTR）¹⁴が立ち上がり、2006年1月時点でICMJEの了承を得るに至った。また時を同じくして、財団法人日本医療情報センター（Japan Pharmaceutical Information Center: JAPIC）¹⁵も臨床試験情報の登録と開示に関する情報提供環境を整備し、広く一般にその情報を公開し2005年7月より運用を開始している。さらに、日本医師会活験促進センター¹⁶において、医師主導活験にかかわる情報が提供されている。

いずれも登録は任意であり、開示する内容についても、登録者側に委ねられている。なおその際、治験¹⁷の登録において、治験薬コード等を含んだ情報を登録・公開した場合、薬事法上の未承認薬

の広告規制に抵触するのではないかとの議論も生じたが、厚生労働省は2005年6月、「直ちに広告には該当しない」旨の見解を示した。これら一連の背景を踏まえ、日本製薬工業協会医薬品評価委員会では「臨床試験の登録・結果公開に関する実施要領」を定めるなど、国内製薬企業においても具体的な運用が始まっている¹⁸。因みに、このような臨床試験登録については、CCとの協力のもとにGlaxoWellcome社が1998年に先陣を切って始めていたことは特筆に価する (Sykes 1998)¹⁹。

エビデンスを正しく伝えるためには、研究の独立性や中立性を積極的に示すことも必要である。例えば、喫煙に関する研究に際し、たばこ産業から支援を受けたか否かを明らかにすることは必要であろう。また特定の団体や企業から支援を受けている研究者に対しても同様で、その団体や企業から支援を受けていたか否かについて公表することにより、他の報告書も含め、研究結果に潜在するバイアスの可能性の有無についての情報を補足することになる。国内では2005年に施行された個人情報保護法との兼ね合いで容易に実現できない可能性もあるが、既にこうした試みは、米食品医薬品局 (Food and Drug Administration: FDA) においては実施されている。

(3) 研究デザインの基礎、ランダム化比較試験 (RCT)

紙面の関係で、各種研究デザインの詳細については他の書籍に譲るが、ここではRCTの要点をレビューするとともに、RCTが保健医療領域でここまで普及した背景について、またその他の領域であまり普及しなかった点について考察する。

上述のバイアスを避ける唯一の方法がRCTのデザインであり、介入 (intervention) の真の効果や因果の推論を行ううえで極めて重要である。仮に未知のバイアスが入り込んだとしても、ランダム化により、そのバイアスを均等に振り分けることができ、集団として同質の背景 (性質) をもつグループ分けが可能となる。この同質の集団に対して、特定の介入を施しその結果を比較することにより、介入の効果のみを純粹に抽出できるところがRCTの研究デザインの本質である。自然変動や

想定しない内外の影響があるため、通常、単純な前後比較から、介入効果のみを抽出することは困難である。

ときに、RCTを「無作為抽出試験」と誤訳されることがある。これは想定する母集団から無作為 (ランダム) に被験者を抽出するプロセス「無作為抽出」 (random sampling) と、抽出された被験者を二群にランダムに振り分けるプロセス「無作為割付」 (random allocation) が混同されたことによる。臨床試験は、抽出したサンプルから得た結果を母集団に対して一般化 (generalize) しようという試みであり、その際、前者の無作為抽出は試験結果の「外的妥当性」を、後者の「無作為割付」は「内的妥当性」を保障するものである。

臨床試験の場でRCTというと通常ブラインド化 (blinding²⁰またはマスキングmasking) が前提となる。Concealment (隠蔽)²¹とマスキングとはよく混同されるが、前者は割付方法や割付順を、介入を行う前に介入を与える人に知られないようにすることである。一方、マスキングは介入が始まった後で患者や医師、結果の評価・解析者などに割付内容を知らせないことである。隠す時点と隠されている対象者が異なる。二重盲検のデザインや中央割付方式によるランダム割付が実施されている場合には、通常concealmentも確保されていると理解できる。

マスキングしない状態で数種類の薬剤を割り付けることを考えると、例えば、現場の医師であれば、割付対象となった重症患者を前に低用量群やプラセボ (偽薬) を割付することに倫理的問題を感じ、割付指示どおりの投薬をしないという行動 (バイアス) が生じる。このような非ブラインド、あるいは単ブラインドの研究デザインの問題点は臨床研究を重ねるうちに経験的に把握されたもので、現在保健医療領域においては、それらの方法から質の高いエビデンスを得にくいとの認識が一般的である。この領域でRCTといえば暗黙の了解の下、二重マスキング (かつランダム化比較) 試験を示すと理解される²²。二重とは治療を施される患者と治療を施す医師の両者がともに、割り付けられた治療方法の中身を知らないという意味である。さらに、必ずしも一般的ではないが、評価者・統計解析担当者にも治療法をマスクする

ことにより、評価・解析時のバイアスを取り除く三重・四重ブラインドのデザインもある。

臨床試験には、①評価を目的とした、②ヒトを用いて、③意図的に開始される、④科学実験、といった四つの視点がある。ヒトを対象として意図的に行うため、当然倫理的問題が生じる。臨床研究においてはその必要性から、こうした問題に正面から取り組み、実施にあたっての様々な手続きや法制化²³によりこれらを解決してきた。例えば、インフォームド・コンセント (informed consent) の下に自由意志で試験に参加することや、事後の同意撤回は今や常識であり、後で同意を撤回してもその後の治療に不利益を受けないことなどについては、臨床試験にかかわる全ての人で合意形成がなされている。

未知の副作用などの負の作用の可能性もあり、必ずしも新しい治療法が優れているとは限らないのであるが、新薬の方が既存薬よりも優れているという一般認識がある。割付の結果、既存薬と比較対照として用いるプラセボ投与を受けることになる可能性についても、事前に説明し合意を得るが、ときに被験者側からどうしても新薬の治療を受けたいといった希望もありうる。このような場合には、クロスオーバー法²⁴により、二種類の治療を時間的に順番を変えて交互に全員に施すことにより参加者への不公平感を是正することも可能である。通常この方法は臨床試験の初期段階において少数の被験者から多くの情報を引き出すために利用されているが、不公平感の是正という観点から教育分野や国際援助の現場でも応用可能である。

臨床研究で経験した様々な問題は、同じく「人」を対象とする人文・社会科学の研究においても同様に生じ得る。保健医療とそれ以外の領域におい

てRCTとSRが普及しない理由のひとつに、これらの手間のかかる問題を克服してまでRCTによるエビデンスを産出する必要性と合理性が認識されない (他の手法で十分であると考えている) ことが推察できる。また一方で、保健医療でエビデンスの集積が可能となった背景には、健康問題は行政も個人も共に直面する強い興味対象であったことがあげられる。さらに、イギリスNHSの研究開発戦略の追い風も受け、適時に予算確保が可能であったことは保健医療領域でのエビデンスの蓄積に大きく貢献したものである。保健医療以外の (場合によってはむしろ予算削減された) 領域は、これらの点で結果的に大きく水を開けられることになったものと考えられる。

なお、保健医療領域外のRCTについては、実施上の制約などからマスキングしない (できない) 状態で試験を実施することの方が多くようである。RCTの品質を読者が評価する場合にも、評価がマスキングされた状態でなされたか、あるいはオープンな状態でなされたか、に依存する系統的な相違が存在することが報告されており (McNutt 1990)、今後RCTの経験が増えるにつれ、保健医療領域で経験した同様の問題が浮上する可能性はある。なお、遺伝情報など個人の利害に関わる情報を扱う場合には、倫理的な観点からその個人が特定できない配慮を施すことは当然であるが、マスキングと患者情報の秘匿とは独立した異なる問題である。

(4) エビデンスの質とお勧め度のグレーディング

エビデンスとは、「バイアスのない方法により得られたデータを、バイアスのない方法で分析して得られた結果」の総称であり、介入結果のほか、

表3 エビデンスのレベルと内容

I a	ランダム化比較試験のメタ・アナリシスによる
I b	少なくとも一つのランダム化比較試験による
II a	少なくとも一つのよくデザインされた非ランダム化比較試験による
II b	少なくとも一つのほかのタイプのよくデザインされた準実験的研究による
III	比較試験や相関研究、ケース・コントロール研究など、よくデザインされた非実験的、記述的研究による
IV	専門家委員会の報告や意見、あるいは権威者の経験

(出所) Agency for Health Care Policy and Research (AHCPR), 1993 (現、Agency for Healthcare Research and Quality)

診断や予後に関わる結果についても産出され得る。そのレベルは科学的研究結果の質や水準に依存し、その研究デザインによって、通常、メタ・アナリシス>ランダム化比較試験 (RCT) >比較臨床試験 (CCT) >コホート研究>症例対照試験 (case-control study) >症例集積研究 (case series) >1例報告 (case report) の順番にレベルが高いとされる。以下に示す、米国の医療政策研究局 (Agency for Health Care Policy and Research: AHCPR, 1993) のグレーディング・スケールもこのような考えによっている (表3)。このスケールの基本的な考え方はその後多くのバリエーションを生んだ。その代表的なものがOxford Centre for Evidence-based Medicine²¹の分類 (2001) として取り入れられている。

EBMにおいては通常、RCT以上のレベルの臨床情報をエビデンスが高いものとして扱うが、非ランダム化あるいはその他の研究スタイルであっても、よくデザインされた研究の結果は、専門家や権威者の報告や意見よりも高いレベルのエビデンスとしている点に注目すべきである。別の見方をすれば、領域により、RCTの実施が困難な場合であっても、よくデザインされた研究を実施すれば当該領域内においてエビデンスを「つくる」ことが可能であることを示している。重要なことは領域毎に異なるエビデンスのレベルを提供者と利用者間で正確に共有することであろう。このようなレベルの概念を含むエビデンスを精確に表現し伝える努力が必要であるが、最終的にはどうしても「つかう」側の認識の問題が介在するため、うまく「つたえる」ための方策や施策が重要となる。

なお、上述したように複数のグレーディング・スケールが並存する中でその運用についての欠点と問題点を改善する目的で有志が集まり、2000年よりGRADE Working Groupとして非公式な活動を開始している²⁶。彼らは、2004年にはエビデンス利用者の視点に立ち、「エビデンスの質」とそのお勧め度 (strength of recommendation) を評価するシステムを開発した (GRADE Working Group 2004, 津谷他訳 2005²⁷)。その中で、個々の主要なアウトカムにおけるエビデンスの質は、(1) 研究デザイン、(2) 研究の質、(3) 一貫性、(4)

直接性、をそれぞれ考慮したうえで決定すべきとしている。そこでは、まず研究デザインをRCTと観察研究 (コホート研究、ケース・コントロール研究、断続的な時系列研究、コントロールされた前後比較研究) に分類し、その介入による利益が害より多いか否かを利害得失の4段階 (正味の利益あり、利害得失あり、不確かな利害得失あり、正味の利益なし) に分類することを提案した。さらに、これらをもとに「お勧め度」のグレード4段階、しなさい (Do it)、した方がよい (Probably do it)、しない方がよい (Probably don't do it)、するな (Don't do it) のカテゴリーを提案した。

彼らの活動の重要な点のひとつは、意思決定に必要な簡潔性 (simplicity) と明瞭性 (clarity) をうまくバランスするよう試みた点である。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り、逆に、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となりやすい。もうひとつの重要な点は、彼らのアプローチは、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行う助けとなるが、ユーザー自身の判断の必要そのものをなくしてしまうものではないとする立場を明確にし評価システムを開発・提案している点である。

(5) 構造化抄録のすすめ

エビデンスを効率的に「つたえる」ための一つの方策として構造化抄録がある。これは学術雑誌編集者らにより提案されたもので、介入研究の場合に限られるがRCTとメタ・アナリシスの論文・報告書の書き方については1996年にCONSORT声明^{28,29}としてまとめられ、国内でも同様の構造を用いた記述が推奨されている (青木2000)。中山らは欧文医学雑誌の構造化抄録書式について検討し、従来から一般に科学論文で用いられてきた抄録記述のIMRAD構造の書式 (Introduction, Method, Results And Discussionの頭文字) と実際に用いられていた8項目の対応を示した (Nakayama 2005)。(1) ObjectiveはIntroduction、(2) Design、

(3) Setting、(4) Patients or participants、(5) Interventions、(6) Main outcome measurementの5項目はMethod、(7) ResultsはResults、(8) ConclusionsはDiscussionに対応したと報告している。

ここで重要な点は、①で研究の目的を明確に記述すること、追従する研究者らが追試を再現できるよう方法については②～⑥を漏れなく書くことである。また、⑦の結果をもとにした単なる議論よりも、⑧で、良かったか、悪かったか、どちらでもないか、といった結論をきちんと書くことが重要である。これらが押さえられていれば、読者は論文の本文を読む前に、研究結果とその結論を容易に知ることができ、その上で必要に応じて試験デザインや目的を知るといった、効率的な情報共有が可能となる。また研究報告者側としても、これらの抄録を書くために、研究開始前から試験目的やデザインを明確にしたうえで、どういう結論を述べるかを事前に十分検討する必要が生じることになり好ましい。さらに、論文化する時点においても標準的な基本構成に従って記述することから、重要な情報の記載漏れを防ぐことにも役立つ。

研究論文の報告形式として構造化抄録が世界的に広く認識されつつある一方で、国内学術雑誌の取り組みが遅れていることも明らかにされた^{30,31,32}。これは、EBMに基づくインフラが国内においてはまだ整備できておらず「つたえる」ことの重要性が十分認識されていないことを示唆する状況である。エビデンスを「つたえる」ところは、コクラン共同計画などの活動だけに委ねるのではなく、エビデンスを生産する研究者も、それを届ける出版社などとともに情報提供環境の整備に貢献できることを再認識すべきであろう。

4. 情報の共有化と社会認識

エビデンスに基づく健康施策や政策を考えると、SRの結果、整理・蓄積された情報を、どのように政策に反映して利用できるかは社会認識 (social epistemology) の視点を含めて考える必要がある。公衆の健康を考える場合蓄積されたエビ

デンスは貴重で有益な情報を与えるが、施策や政策が他の諸々の情報をもとに決定されることを踏まえれば、いかに明白なエビデンスであろうとも、ひとつの要素に過ぎない。エビデンスを政策決定へ統合させ、広く社会に還元させるために、どのような問題と解決策があるか考察した。

(1) 説明のジレンマ

マスメディアによる(エビデンスや保障のない)興味本位の報道の方が容易に社会に伝播・浸透する現実があり、ときに、重要なエビデンス情報を伝えることの弊害にもなり得る。これをうまく回避する方法論が必要になるが、一般に、複雑な情報を専門家でない人に説明する場合、細かく説明すればするほど(情報量が増せば増すほど)かえって本質の理解から遠ざかるということはしばしば経験される。いわば「説明のジレンマ」とも言うべき障壁である。個人の理解を促す最適な情報量というものは伝達する情報毎に必ずどこかに存在するはずであるが、それは個人の背景知識や認知能力にも依存し、また、質的にも全く異なるものと理解すべきである。従って、「説明のジレンマ」を前に、個人に適切かつ正確に要点を伝える共通の道具や方法論にはあまり期待がもてない。

一方、個人に対してエビデンスが正確に伝わらない場合でも、社会全般にエビデンスから得られた施策や思考パターンなどを流布させることは可能である。「説明のジレンマ」が生じにくい(と考えられる)レベルの情報量を比較的短時間で繰り返し伝達することにより、形骸的な部分は容易に伝播する。ラジオ・テレビのコマーシャルや広告チラシによる、日常生活における知識の伝播がそれである。斬新さはないが、既に一般に定着しているメディアであり、情報の受け手の認識力のトレーニングや特別なスキルは必要ない。しかしこの方法は、一つの現実的な対応策であるが、制御不能の過剰反応を生じる危険性も伴う。情報の受け手の幅が極めて広いため、エビデンスから得られた本質的な情報を正確に伝達できない可能性もあり、ときに情報内容が意図せず誇張されたり、伝達内容が多様化し全く異なる方向に認識されるという問題も生じ得る。

同じ情報を流した場合に生じるこれらの問題は、受け手側の認識上の問題であるが、その原因は多種多様であり全体としての解決方法はなさそうである。

コミュニケーションの問題は、意思を伝えようとする側とその受け手が存在する限り、常に不確定な要素が介在するため、公衆を対象とした優れたキャッチ・フレーズがいつも見つかるとは限らない。

(2) 知識伝播と社会認識

科学的な意味やその背景が伝わらなくとも、特定の機会を通じて、エビデンスの本質を社会に伝達・認識させることは可能である。在住のヘルス・ケア専門家の協力を得ることにより、地域のコミュニティーや学校教育の場を通じて、場に適した情報を効率よく提供し共有できる。直接顔を合わせた双方向性の指導は、一般を対象とした一方向性の情報よりも、大きな影響力を持つことは経験的に明らかにされている (Redelmeier 1993)。

また、日常診療の変更に関して、地域の専門家による勧告は国が公表したエビデンスに基づく一般的なガイドラインよりも強い影響力を持つことも報告されている (Lomas 1991)。このような背景を考えると、上述のマスメディアを用いた、大まかな知識伝達を、こうした地域レベルのきめ細かなフォローアップにより支援することが最も効率よく社会認識を醸成できる方法のように思われる。開発途上国における筆者らのヘルス・プロモーションの経験からも、このような国レベルのマクロなキャンペーンと地域レベルの実践的なフォローアップによるアプローチが、社会認識に有効であることは十分理解できる。

エビデンスの知識伝播と社会認識において、もうひとつ課題がある。エビデンスを効果的に伝えるコンテンツの作成がそれである。ほとんどの保健医療に関わる論文は、研究者により他の研究者のために書かれたものである。これを利用者によりわかりやすく伝える方法が必要である。すなわち、伝統的な紙媒体の、また専門用語を用いて精緻に記述された論文は、恐らくほとんどの人にとっては全くつまらない読み物であり、明確にエビデ

スが記述してあったとしてもほとんど意味をなさない。このような場合には、セールス、初等教育、グラフィックデザインや広報を専門とする人々の知識や技術を学び協力する必要がある (ハダット 2004, p.118)。特に、次世代の担い手らの特質も踏まえて、適切なメディアを利用することが肝要である。初等教育においてエビデンスについて「つたえる」ことは一見困難かのように思えるが、適切なメディアがあればその本質部分をうまく伝えることは可能と思われる。

社会認識が伴わないために、多大のリソースをつぎ込んで構築したエビデンスの宝庫が公益に資さないとすれば、これは深刻な問題である。上述のいずれの方法によるにしても、その基本として、科学の公衆的理解 (public understanding of science) の振興が鍵となる。このあたりの方策と議論については、雑誌「臨床評価」の特集「ナレッジ・マネジメントと情報の共有化」に詳しい²⁹。また、集団を対象とするEBHにナレッジ・マネジメントの概念等を導入し、エビデンスに基づく意思決定、および研究から実践へのアプローチについてはミュア・グレイの著書 (2版) が参考となる (ミュア・グレイ 2005³⁴)。この意思決定や研究から実践へのアプローチの重要性と、同書で述べられている「正しいことを正しく行う (Doing the right things right)」ことの意味するところは深い。

また、エビデンスを適切に伝播し社会認識を喚起するための評価コミュニケーションは、健康問題をはじめとするあらゆる領域におけるリスク・マネジメントの視点からも重要である。さらに、EBMの方法論と問題点、現代科学論、科学哲学の視点からの考察は芥尾らの総説が参考になる³⁵。

(3) 量か質か？

RCTやSRの話をすると必ず昔ながらの論争が噴出する。量的研究 (RCTとSRにも応用される統計学的手法の適用が可能な研究) 対、質的研究 (個人や社会の内面の定性的な側面を対象とする研究) の対峙、いわゆる研究パラダイム (対象の把握の仕方) 論争である。

RCTやエビデンスに纏わる多くの論争については、量か質かのどちらかに偏った保守的な考え方

に加え、単に、互に各々の評価研究法を未経験であることによる不理解があるように思われる。あるいは、量的研究においては数学や数理統計が共通言語と成り得るが、質的研究にはこれに対応する共通言語がないことがコミュニケーションを難しくしている原因かもしれない。

様々な議論は省略するが、明白にいえることは、RCTやSRはむしろ定性的・質的な研究を支援するものであり、あらゆる既存研究の成果を相補的に高めるものであって、決して既存研究と齟齬きたすものではないということである。必ずしも一般的ではないが、医学界に強調されてきた「根拠」や「統計」「科学性」に対する補完的な意味をもつ考えとしてNBM (Narrative-based Medicine) も提唱されている (グリーンハル2001)。質的研究法により量的研究の欠陥が明らかになることもあるので、量的研究と質的研究両者のバランスよい協調は本来の目的 (対象から得られる真実を明らかにすること) への近道である。

1993年のPopeらの論文 (質的研究をしている社会学者と保健医療サービス研究部門の責任者の間で交わされた架空の議論を哲学者らの問答のようにした内容) を契機に、保健医療の分野における質的研究についての注目が高まり、British Medical Journal誌の連載を経て1996年に単行本が出版された。この世界的に著名な医学誌に連載が始まってから、イギリスでは保健医療分野で質的研究が盛んに行われるようになった。質的研究が様々な問題を理解することに役立つことが注目され、2000年には改訂2版も上梓され、さらに内容を充実させている。質的研究においても研究の質の問題が生じ、質の「良いもの」と「悪いもの」を区別する方法を模索することとなった。英国NHS Research and Development Programmeは、1998年頃医療技術評価に関する質的研究手法を再検討するための研究費を付けたが、残念ながら現時点でまだガイドラインは出ていないようである。

このように最近の研究上の疑問に答えるためのより一貫した効果的な方策により、定量的アプローチと定性的アプローチが併合される方向性が模索されている (Pope 1995)。2005年10月、メルボルンで開催された第13回Cochrane Colloquiumにお

いて「エビデンスは既に十分か?」との疑問のもとに議論が行われ、得られたエビデンスは単独ではなく、必要とされる場で実際に利用可能な形で研究の成果と統合されるべきであるとの再確認がなされた (Summerskill 2005)。Web検索でこの方面の動向を眺めると英国Leicester大学のQual-Quan Evidence Synthesis Groupなどがこの革新的な境界領域の研究を進めていることがわかる³⁶。ここでは、蓄積したEBHにおけるSRの経験をもとに、質的研究論文のSRを行い、質的エビデンスを併合する研究を行っている。また、2006年6月末、EBH推進の担い手となっているOxford大学のCentre for Continuing Professional Developmentは、Qualitative Research and Systematic Reviews Workshopを開催され³⁷、コクラン共同計画のCochrane Qualitative Methods Review Groupに経験のフィードバックを行うようである。EBM/EBHの系譜を踏まえ、今後の研究動向に注目したい。

おわりに

わが国における人文社会科学領域におけるエビデンスの産出とその情報基盤の整備は、次世代に向けての我々の責務である。歴史的考察から得た教訓として、政府は明確な戦略を掲げたくて、ガイドラインや法的な整備を進めるとともに、適切な予算確保の仕組みを構築する必要がある。また同時に、評価研究を行うわれわれは諸外国の評価研究者や既に活動を推進しているグループとの連携を深める必要がある。各々の役割と専門性を活かし、ひとつひとつ、社会認識を醸成するための布石を置き続けてゆくことが肝要である。澎湃たる未来は、次世代に向けた長期的視野に立つ施策と個人の科学理解および社会認識能力の高揚にゆだねられている。

注記

- 1 EBMにおけるevidence は、「バイアスのない方法により得られたデータを、バイアスのない方法で分析して得られた結果」の総称である。介入結果のみならず診断や予後についても産出され、その産出の方

- 法や研究の質によりそのレベルは異なる。「証明力」「証拠」「科学的根拠」「科学的事実」などと訳されることがあるが現時点で定訳はない。本稿では「エビデンス」を用い、一般的用語のevidenceと区別した。
- 2 これまでに系統的な調査は知られていないが、各分野の専門家らとの情報交換によりこの状況は共通の課題として認識される。筆者らの調査からもこの状況が確認されている。
 - (1) 津谷喜一郎、津富宏、正木朋也 (2003) 「人文社会科学分野におけるランダム化比較試験などの研究デザイン使用の現状」、『厚生労働科学研究費補助金 (医療技術評価総合研究事業) 「日本におけるEBMのためのデータベース構築及び提供利用に関する調査研究 (H14-医療-067) 平成14年度 総括・分担研究報告書」 (主任研究者：丹後俊郎)、183-189
 - (2) 正木朋也、津富宏、三砂ちづる、加藤則子 (2003) 「青少年暴力に関わる国内研究動向の文献的検討」、『厚生労働科学研究費補助金 (こころの健康科学研究事業) 「思春期における暴力行為の原因究明と対策に関する研究 平成14年度研究報告書」 (主任研究者：小林秀資)』、85-98
 - 3 本部のホームページは英語を含め7ヶ国語で閲覧可能である (<http://cochrane.org/>)。一方、国内ではJANCOC (Japanese informal Network for COchrane Collaboration) に情報が集約されている (<http://cochrane.umin.ac.jp/>)。
 - 4 本稿では狭義のEBMには殆ど触れていない。医療現場の実践としてのEBMについては多数の書籍が出版されているので、それらを参照されたい。
 - 5 原本 J.A.Muir Gray (1996). Evidence-Based Healthcare - How to Make Health Policy and Management Decisions. Churchill Livingstone. なお、本書の日本語訳は1999年に発行されている。また、2001年に第二版が上梓され、その日本語訳も2005年に出版されている。
 - 6 5のMuir Grayの本の表現はhealthcareが一語となっている。この他、health care、あるいは、HealthCareとしている場合も見受けられる。
 - 7 ファイヨール著、山本安次郎訳 (1985) 『産業ならびに一般の管理』、ダイヤモンド社
アンリ・フェイヨール著、都筑栄訳 (1964) 『産業並に一般の管理』、風間書房
 - 8 John Wiley and Sons, Inc (Cochrane Library Customer Services Advisor Alison Robb on: +44 (0) 1243 843379, e-mail: arobb@wiley.co.uk) に直接申し込むほか、日本では、南江堂洋書部 (Tel: 03-3811-9950, Fax: 03-3811-5031, e-mail: nkdyosho@nankodo.co.jp) が代理店となっている。
 - 9 Centre for Reviews & Dissemination, University of York, UK, (<http://www.york.ac.uk/inst/crd/>)
 - 10 日米欧三極のハーモナイゼーション会議 The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH: <http://www.ich.org/>) が作成したガイドラインのひとつ。日英対訳および関連情報は以下に詳しい。 <http://home.att.ne.jp/red/akihiro/index.htm>
 - 11 原本: Alejandro R. Jadad (1998). Randomized Controlled Trials: A User's Guide. BMJ Books.
 - 12 Journal of Negative Results in Biomedicine, <http://www.jnrnm.com/>
 - 13 International Clinical Trials Registry Platform (ICTRP) (<http://www.who.int/ictip/>)
 - 14 UMIN臨床試験登録システム (<http://www.umin.ac.jp/ctr/>)
 - 15 財団法人日本医療情報センター (<http://www.japic.or.jp/>)
 - 16 日本医師会治験促進センター (<http://www.jmacct.med.or.jp/>)
 - 17 「患者や健康な人を対象とした治療を兼ねた試験」が「臨床試験」であり、そのうち特に新薬開発 (医薬品製造承認申請) を目的とした「臨床試験」を「治験」と呼ぶ。
 - 18 臨床試験の登録・結果公開に関する実施要領 2005 <http://www.jpma.or.jp/policy/pdf/rinsyou.pdf>、同 Q&A <http://www.jpma.or.jp/policy/pdf/rinsyouqa.pdf>
 - 19 企業合併にともない一時中断したが現在も以下に引き継がれている。 (<http://ctr.gsk.co.uk/>)
 - 20 盲検 (blinding) という用語は、盲人蔑視の視点が含まれるとして日英ともに使用を控えるべきとの議論もある。代案として「遮蔽」「マスクキング」などが提案されている。
 - 21 著者らは「隠蔽」と訳している。隠蔽と訳される場合もある。
 - 22 臨床研究の現場で、二重盲検ランダム化比較試験という語は使わない。盲検化すればランダム化するこ

- とは通常包含されるため、ランダム化を省略して二重盲検比較試験あるいは、比較も省略し二重盲検試験と呼ばれる。
- 23 日米欧三極のハーモナイゼーション会議 (ICH) での協議を経た「医薬品の臨床試験の実施に関する基準 (Good Clinical Practice: GCP)」や「疫学研究に関する倫理指針」などがある。いずれも、ヒトと対象とする医学研究の倫理的原則であるヘルシンキ宣言にその起源をみることができる。
- 24 クロス・オーバー法: 最も単純な例として、二種類の介入AとBの比較を行う際、ABの順に実施する群とBAの順に実施する二の群を設け、集計解析時にAとBの介入をそれぞれまとめてA群とB群を比較する方法。
- 25 Oxford Centre for Evidence-based Medicine (<http://www.cebm.net/index.asp>)
- 26 GRADE working group (<http://www.gradeworkinggroup.org/>)
- 27 原本: GRADE Working Group (2004). Grading Quality of Evidence and Strength of Recommendations. *BMJ*,328 (7454) :1490-7.
- 28 Begg,C.他 (著)、津谷喜一郎、小島千枝 (訳) (1997) 「無作為化比較試験の報告の質を改善する方法 — CONSORT声明」『*JAMA<日本語版>*』、74-78
- 29 Rennie,D. (著)、津谷喜一郎、小島千枝 (訳) (1997) 「無作為化比較試験をいかに報告するか — CONSORT声明」『*JAMA<日本語版>*』、78-79
- 30 中山健夫 (2001) 「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」、『平成13年度 総括・分担研究報告書』、厚生科学研究費補助金 21世紀型医療開拓推進研究事業 (<http://mhlw-grants.niph.go.jp/>)
- 31 同上、『平成14年度 (2002) 総括・分担研究報告書』
- 32 同上、『平成15年度 (2003) 総括・分担研究報告書』
- 33 臨床評価刊行会 (2002) 「特集: ナレッジ・マネジメントと情報の共有化」、『臨床評価』、29 (2・3) :222-305
- 34 原本: J.A.Muir Gray (2001). Evidence-Based Healthcare - How to Make Health Policy and Management Decisions. 2nd Ed., Churchill Livingstone.
- 35 齊尾武郎、栗原千絵子 (2001) 「Evidence-based medicineの現代科学論的考察 — EBM questioned; a metascientific criticism」『臨床評価』、29 (1) :185-201
- 36 Qual-Quan Evidence Synthesis Group, Biostatistics & Genetic Epidemiology Group, part of the Department of Health Sciences, at the University of Leicester. <http://www.hs.le.ac.uk/research/qualquan/index.htm>
- 37 http://www.conted.ox.ac.uk/cpd/healthsciences/courses/short_courses/qsr/speaker.asp

参考文献

- Feinstein,A.R. (1968). Clinical judgement. Huntington, NY: RE Krieger Pub. Co.
- Guyatt,G.H. (1991). Evidence-based medicine. ACP Journal Club March/April,A-16.
- Jones,R. (1995). Who does qualitative research? *BMJ*,311,2.
- Lomas,J., Enkin,M., Anderson,G.M., Hannah,W.J., and Vayda,E. (1991). Opinion leaders vs audit and feedback to implement practice guidelines. *JAMA*,265,2202-7.
- McNutt,R.A., Evans,A.T., Fletcher,R.H., and Fletcher,S.W. (1990). The effects of blinding on the quality of peer-review. *JAMA*,263,1371-6.
- Nakayama,T., Hirai,N., Yamazaki,S. and Naito,M. (2005). Adoption of structured abstracts by general medical journals and format for a structured abstract. *J Med Libr Assoc*,93 (2) April,237-242.
- Owen,R. (1982) Reader bias. *JAMA*,247,2533-4.
- Paul,J.R. (1938). Clinical epidemiology. *J Clin Invest*,17 (5) September,539-541.
- Pope,C. and Mays,N. (1993). Opening the black box: an encounter in the corridors of health services research. *BMJ*,306,315-8.
- Pope,C. and Mays,N. (1995). Reaching the parts that other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ*,311,42-5.
- Redelmeier,D.A., Rozin,P., and Kahneman,D. (1993). Understanding patient' decisions: cognitive and emotional perspectives. *JAMA*,270,72-6.
- Rossi,P.H., Freeman,H.E. and Lipsay,M. (1999). Evaluation: A Systematic Approach 6th Edition.

- Thousand Oaks, CA:Sage Publication.
- Sackett,D.L., Haynes,R.B., Guyatt,G.H., et al (1991).
Clinical epidemiology - A basic science for clinical
medicine. Second Ed. Little, Boston/Tronto/London:
Brown and Company.
- Sackett,D.L., Richardson,S. Rosenberg,W. and Haynes,R.B.
(1996). Evidence-based medicine: how to practise and
teach EBM. Edinburgh: Churchill Livingstone.
- Smith,R. (1988). The sickening of medical research.
BMJ,296,1079-80.
- Summerskill,W. (2005). Cochrane Collaboration and the
evolution of evidence. Lancet,366,1760 (Comment).
- Sykes,R. (1998). Being a modern pharmaceutical
company. BMJ,317,1172-80. (翻訳 サイケス, リチャ
ード (2000) 「現代的製薬企業であること—それ
には臨床試験の情報公開プログラムを欠かすことが
できない」、『臨床評価』、27 (3) , 509-10
- グリーンハル, トリシャ、ハーウィッツ, プライアン
編/斎藤清二、山本和利、岸本寛史監訳 (2001)
『ナラティブ・ベイスド・メディスン』、金剛出版
- ミュア・グレイ,J.A. (著)、久繁哲徳 (監訳) (1999)
『根拠に基づく保健医療—健康政策と経営管理の判
断決定の方法』、じほう
- ミュア・グレイ,J.A. (著)、津谷喜一郎、高原亮治
(監訳) (2005) 『エビデンスに基づくヘルスケア
ヘルスポリシーとマネージメントの意思決定をどう
行うか』、エルゼビアジャパン
- ハダット、アレハンドロ (著)、舟喜 光一、津谷喜一
郎 (訳) (2004) 『ランダム化比較試験—100の
questionsに学ぶ』、じほう
- ポーブ、キャサリン、メイズ、ニコラス (編)、大滝純
司 (監訳) (2001) 『質的研究実践ガイド』、医学書院
- 青木仕 (2000) 「構造化抄録の基礎知識」、中嶋宏 (監
修)、津谷喜一郎、他 (編) 『EBMのための情報戦
略—エビデンスをつくる、つたえる、つかう—』、
中外医学社、82-93
- 窪田好男 (1998) 「N P M型政策評価と政府の失敗—
地方行革との関連で—」、『会計検査研究』、18:47-62
- 佐々木亮 (著)、龍慶昭 (監修)、(2003) 『政策評価ト
レーニングブック』、多賀出版
- 田辺智子 (2003) 「米国90年代の行政改革」、『レファ
レンス』、30-46
- 津谷喜一郎 (1995) 「情報の質の確保」『臨床と薬物治
療』、14 (7) :593-7
- 津谷喜一郎 (1996) 「コクラン共同計画とシステマテ
ィック・レビュー」、『薬理と治療』25 (1) :11-24
- 津谷喜一郎 (1999) 「EBMとコクラン共同計画」、矢野
栄二 (編) 『医療と保健における評価—Toward
Evidence Based Medicine—』、南江堂、195-217
- 津谷喜一郎 (2000a) 「コクラン共同計画とは — 日本
的展開へ向けて」、『公衆衛生研究』、49 (4) :313-9
- 津谷喜一郎 (2000b) 「コクラン共同計画とシステマテ
ィック・レビュー—EBMにおける位置付け—」、厚
生省健康政策局研究開発振興課医療技術情報室 (監
修)、『わかりやすいEBM講座』 厚生科学研究所、
120-35
- 津谷喜一郎 (2000c) 「医学研究デザインの基礎」、中
嶋宏 (監修)、津谷喜一郎、他 (編) 『EBMのための
情報戦略—エビデンスをつくる、つたえる、つかう
—』、中外医学社、22-47
- 津谷喜一郎、中山健夫、島村治子 (訳) (2005) 「エビ
デンスの質とお勧め度のグレーディング」『薬理と
治療』33 (12) :1241-54
- 林謙治 (2000) 「根拠に基づく健康政策へのアプロ
ーチ」、『公衆衛生研究』、49 (4) :346-53
- 龍慶昭、佐々木亮 (著) (2000) 『「政策評価」の理論
と技法』、多賀出版

(2006.2.28受理)