

解 説

## Evidence-Based Medicine (EBM) 入門ガイド (4)

## 治療に関する文献の評価

<sup>1</sup> 東京女子医科大学 衛生学公衆衛生学 (II)<sup>2</sup> 同 衛生学公衆衛生学 (I)<sup>3</sup> 同 内分泌疾患総合医療センター 外科コジマハラノリコ・カガワ ジュン オカモト タカヒロ  
小島原典子<sup>1</sup>・香川 順<sup>2</sup>・岡本 高宏<sup>1,3</sup>

(受付 平成11年10月29日)

## Guide for Evidence-Based Medicine (4)

## Approach to the Evaluation of Medical Treatment

Noriko KOJIMAHARA<sup>1</sup>, Jun KAGAWA<sup>2</sup> and Takahiro OKAMOTO<sup>1,3</sup><sup>1</sup>Department of Hygiene and Public Health II<sup>2</sup>Department of Hygiene and Public Health I<sup>3</sup>Department of Endocrine Surgery, Institute of Clinical Endocrinology  
Tokyo Women's Medical University

We often hear about evidence-based medicine (EBM) in medical journals or speeches at conferences. When you read about EBM and try to practice it, you may encounter some difficult terms such as “critical appraisal” or “relative risk”. When you see statistical expressions, many of you may ever give up. In the 1999 clinical clerkship for 6th year medical students, we make sure you can acquire the skills in technical concepts and terminology defined through work sheets, which we have reported and associated with the Users' Guide series published in the Journal of the American Medical Association. By using these sheets, you can systematically check through the validity, importance, and ways to adapt them to your actual patient. As a definite plan, we chose the WOS (West of Scotland) study conducted by Shepherd et al well-known as the primary prevention study for hypercholesterolemia. At first, for the evaluation of validity, you should check the following: if the study is a randomized controlled trial, how were endpoints defined? Second, you should evaluate the importance of the results using relative risk (RR), relative risk reduction (RRR) and the number needed to treat (NNT). Finally you should investigate how your actual patient can be adapted to the process. It will be an important skill for clinicians in the future to acquire the skill of EBM. However, we recognize that the well being of patients supersedes any recommendations we make through EBM. Still, it is necessary for medical experts to continue to look for evidence that does not rely solely on our experience or impression.

## はじめに

最近, evidence-based medicine (EBM) とか臨床疫学という言葉が雑誌のタイトルや講演会などで耳にする機会が増えた。少し興味をもって本を読んでみると批判的吟味や相対危険度などの難しい言葉が並び, 更に統計学の難しい数式がでてきて諦める, というパターンが現状では多いのではないだろうか。我々は, EBM を日常診療に導入する方法として JAMA の Users' Guide シリーズ<sup>1)~3)</sup>をもとにしたワークシートを提唱してきた<sup>4)</sup>。1999 年の東京女子医科大学医学部 6 年生に対するクリニカル・クラークシップに EBM を導入した経験からも, 具体例をワークシートに書き込む作業を通して, 言葉の意味, 値の重要性を体験した方が理解が深まると考えている。

例えば, 総コレステロール 258mg/dl の患者さんが来院したとき, 「総コレステロール値でみると正常の人に比べて将来の虚血性心疾患の発生率 (相対危険度) が 2 倍以上になるので薬を飲む方がよい」と判断できる臨床医は何割ぐらいいるだろうか。高コレステロール血症のために動脈硬化が進む可能性があるということがわかっている, 具体的に危険の度合を知らなければ, その患者の治療の重要性が判断できず, 結局患者に対する説得力も弱くなってしまふ。そこで, 情報を検索し批判的に吟味するわけである。このように問題を明確にし, 根拠に基づいて合理的に行う医療が EBM である。

高脂血症の大規模臨床試験 (メガスタディ) は, 質の高いものが多く治療の分野の EBM の体得には最適と思われる。情報検索は他稿に譲り, 今回は, Shepherd ら<sup>5)</sup>による WOS (West of Scotland) 研究の論文を実際に批判的に読み, EBM に必要な言葉の意味や統計学的な値を学習してみる。ワークシートに書き込むことで論文の妥当性, 重要性が検証でき, 実際に自分の患者に当てはめられるか検討できるようになっている。

本稿では“治療に関する EBM”に焦点を絞っているが, 本連載を通して体系的な EBM の理解の助けになれば幸いである。

## 治療に関する批判的吟味

### 1. シナリオ

生来健康な 58 歳の女性が人間ドックを受診して検査を受け, 総コレステロール値が 258mg/dl であった。彼女はあなたの外来を受診し, 治療が必要かどうかを尋ねた。

### 2. 情報収集

#### 1) 疑問の明確化

- ①患者 patient: 生来健康な 58 歳の日本人女性
- ②介入 intervention: pravastatin
- ③比較 comparison: 偽薬 placebo
- ④結果 outcome: 虚血性心疾患

#### 2) 情報検索

- ①原著論文データベース: MEDLINE
- ②検索のためのキーワード: hypercholesterolemia AND pravastatin

EBM の実践は自分が何を知りたいのかを明確にすることから始まる (疑問の明確化)。そのためには疑問の 4 要素をまず明らかにする。患者はこれまでに虚血性心疾患の既往のない 58 歳の日本人女性である。介入 (治療) として現在広く使われている pravastatin を考えよう。そして pravastatin と比較するのは偽薬 placebo である。治療の有効性を判定する結果項目 outcome として虚血性心疾患の発生率を使うことにする。ただし, 何が最重要の結果項目であるかは臨床家の判断だけでなく, 患者の価値観などによっても異なる場合があることに留意すべきである。

これら 4 要素が明確になったら, 自分が知りたい疑問を一つの文で言い表してみる。「総コレステロール値が 258mg/dl の生来健康な 58 歳の日本人女性で pravastatin による治療は偽薬と比べて虚血性心疾患の発生率を減少させるか?」

この疑問に対して情報の収集を行う。情報検索は非常に大事なステップであるが, 阿字川ら<sup>6)</sup>がすでに解説を行っているので今回は割愛する。Shepherd ら<sup>5)</sup>による WOS 研究の論文を用い, 我々の使用しているワークシートの具体例を表に示す。これにそって治療に関する文献の吟味の 1 例を提示する。

### 3. 批判的吟味

#### 1) 研究結果の妥当性 (validity)

(1) 研究のデザイン：治療効果の検証に最も妥当な研究デザインはランダム化比較試験 (randomized controlled trials; RCT) である。

(2) 対象集団の合致基準と除外基準：臨床試験の対象となった人々の特徴を知ることが、結果を目の前の患者に当てはめる際に必要なことである。具体的には、疾患の定義と程度、対象者の年齢、性別、人種などを念頭におく必要がある。

(3) 比較する治療 (介入) の内容を列挙して明確にする。

(4) 臨床的に重要な結果 outcome の調査 (測定) 法が明確に示されているか、その方法の信頼性 (再現性) と妥当性 (正確さ) が記載されているか? : エンドポイントとして何を指標にしているか、どんな評価法で測定しているかも重要である。死亡率で評価してある場合はばらつきは少ないが、疾患発生率で評価する場合は、診断基準、検査方法によってばらつくことが多い。また、どの程度まで追跡調査をしているか、例えば、死亡診断書の病名だけなのか、カルテを取り寄せて検討しているかによっても結果が異なる可能性がある。

(5) Blinding: 被検者や治療効果を判定する者

表 EBM「治療の批判的吟味ワークシート」

筆頭著者名	James Shepherd et al		
タイトル	Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia		
掲載誌名	N Engl J Med 1995, 333(20) : 1301-1307		
＜研究結果の妥当性＞			
1.	Design	ランダム化試験であるか?	<input checked="" type="checkbox"/> はい <input type="checkbox"/> いいえ
2.	Population	8週間の食事療法後、血清コレステロール値 272 ± 23mg/dl 45～64歳の男性患者 6,595人 心電図異常、虚血性心疾患の既往のあるものは除外	
3.	Intervention	介入は何か?	(1) pravastatin 40mg 1×夕 (2) placebo
4.	測定	最重要の結果の測定法について信頼性と妥当性の記載はあるか?	<input checked="" type="checkbox"/> 記載あり <input type="checkbox"/> 記載なし
5.	Blinding	治療を受ける者と効果を評価する者は、治療内容を知っていたか?	被験者 <input type="checkbox"/> はい <input checked="" type="checkbox"/> いいえ 評価者 <input type="checkbox"/> はい <input checked="" type="checkbox"/> いいえ
6.	治療開始時の背景因子	効果に影響を及ぼしうる要因について、治療群間で差がなかったか?	<input checked="" type="checkbox"/> 差はなかった <input type="checkbox"/> 差はあった <input type="checkbox"/> 記載なし
7.	Contamination	割り付けられたものとは別の治療を誤って受けた被験者がいなかったか?	<input checked="" type="checkbox"/> いなかった <input type="checkbox"/> いた <input type="checkbox"/> 記載なし
8.	Co-intervention	検証した治療法に加えて、その他の治療も受けていた被験者がいなかったか?	<input checked="" type="checkbox"/> いなかった <input type="checkbox"/> いた <input type="checkbox"/> 記載なし
9.	Follow-up	追跡数 / 開始時例数 治療(1)群 <u>正確な記載はなし</u> 治療(2)群 <u>正確な記載はなし</u>	追跡期間：平均 4.9年 追跡率：約 30%
10.	脱落例の扱い	データ分析に際しての脱落例の扱い	<input checked="" type="checkbox"/> 含めて分析 (ITT) <input type="checkbox"/> 除外して分析 <input type="checkbox"/> 記載なし

つづく

## &lt;結果の重要性&gt;

		介 入	
		治療群	コントロール群
結果	イベントあり	a 174	b 248
	イベントなし	c 3,128	d 3,045
	計	a + c 3,302	b + d 3,293

  

		治療効果(95% 信頼区間)
1.	相対リスク減少 RRR = $1 - \{a/(a+c)\} / \{b/(b+d)\}$	0.30(0.16 ~ 0.42)
2.	絶対リスク減少 ARR = $b/(b+d) - a/(a+c)$	0.022(0.011 ~ 0.034)
3.	NNT = $1/ARR$	45(29 ~ 92)
4.	点推定値の重要性	<input checked="" type="checkbox"/> 臨床的に重要である <input type="checkbox"/> 臨床的に重要でない <input type="checkbox"/> わからない
5.	区間推定の重要性	<input checked="" type="checkbox"/> 統計的に有意である <input type="checkbox"/> 統計的に有意でない

## &lt;研究結果の適用可能性&gt;

1.	この研究結果の population は、あなたの患者背景と同じか(年齢層、性別、人種、疾患の重症度、など)?	<input type="checkbox"/> はい <input checked="" type="checkbox"/> いいえ <input type="checkbox"/> わからない
2.	あなたの患者はその治療効果に満足するだろうか?	<input type="checkbox"/> はい <input type="checkbox"/> いいえ <input checked="" type="checkbox"/> わからない
3.	あなたの患者は治療に伴う費用や危険の程度と頻度を受け入れるだろうか?	<input type="checkbox"/> はい <input type="checkbox"/> いいえ <input checked="" type="checkbox"/> わからない

が治療の割り付けを知っていると、結果に重大な影響を及ぼす可能性がある。例えば、死亡原因を判定する際に pravastatin を投与されていたことを知っていたら、判定者の診断が偏る可能性がある。それを避けるためには、彼らが治療内容を知らない状態にすることが望ましい。これを blinding という。

(6) 治療開始時の群間同等性：俗に「背景因子」と呼ばれているが、治療開始に当たって重要な因子について群間に差がなかったかどうかを吟味する。高脂血症の治療であれば性別、年齢の他、虚血性心疾患の危険因子（高血圧や糖尿病の有無、喫煙歴など）で群間差がなかったかどうか重要であろう。

(7) Contamination：敢えて和訳すれば「誤分

類」である。割り付けられたものとは別の治療を誤って受けることをいう。例えば、pravastatin を割り付けられたのにプラセボを投与されてしまうなどである。

(8) Co-intervention：臨床試験で検証しようとしている治療法以外の介入を受けることをいう。例えば被検者が高脂血症に非常に関心があり、薬物療法（pravastatin またはプラセボ）に加えて食事療法や運動療法を熱心に行えば、研究結果の大きさは真実から偏る可能性がある。

(9) Follow-up：追跡できずに最終結果（転帰）がわからない被検者は、分析から除外せざるをえない。しかし、除外すると無作為化による群間同等性がくずれてしまう。例えば“Evidence-Based Medicine”などの雑誌は追跡率 80% 未満

の論文は採用しないという。

(10) 脱落例の取り扱い：治療を継続できず脱落しても最終結果（転帰）がわかっているなら、割り付けられた治療群に含めて分析する。これを intention to treat analysis (ITT) という。ITT を結果分析の原則とする理由は二つある。第一にこれを守ること無作為化による群間同等性を保つことができる。第二に「治療を継続できない」症例に遭遇することは臨床の現場ではめずらしくないので分析結果が「より実践的である」といえる。

## 2) 研究結果の重要性 (importance)

(1) 治療効果の点推定 (臨床的重要性)：治療効果の指標には以下のものがある。これらの指標による治療効果の大きさが臨床的に重要なものであるかどうかの判断には、臨床家の専門的な知識や経験を要する。

① 相対危険減少 (relative risk reduction; RRR)：リスクの相対的な大きさを表すのが相対危険度 (RR) であり、介入群の罹患率を  $X_e/N_e$ 、非介入群の罹患率を  $X_u/N_u$  とすると  $RR = (X_e/N_e)/(X_u/N_u)$  で表わされる。これが1より小さければ、罹患の危険が介入により減少したということである。WOS 研究では pravastatin 群での虚血性心疾患発生率が 174/3,302 (5.3%)、偽薬群での発生率が 248/3,293 (7.5%) であった。よって相対危険度は  $5.3/7.5=0.7$  である。そして具体的なリスク減少の表現として相対危険減少は  $RRR=1-RR$  で与えられる。Pravastatin による虚血性心疾患の相対危険減少は  $1-0.7=0.3$  (30%) である。

② 絶対危険減少 (absolute risk reduction; ARR)：リスク減少を介入群と非介入群との罹患率の差で表したものが絶対危険減少 (ARR) である。WOS 研究では pravastatin 群での虚血性心疾患発生率 5.3% と偽薬群での発生率 7.5% の差であり、 $ARR=7.5-5.3=2.2\%$  である。つまり、pravastatin 投与は、偽薬に比し、虚血性心疾患発生率を 2.2% 余計に減少させることができることを示している。

③ NNT (number needed to treat)：有害な結果が1例発生するのを防ぐために治療しなければならない患者数であり、絶対危険減少の逆数で与え

られる。WOS 研究では  $ARR=2.2\%$  であったので  $NNT=1/2.2\%=45$  となる。すなわち1人の虚血性心疾患発生を防ぐために pravastatin を45人に投与する必要があるということになる。

(2) 治療効果の区間推定 (統計的有意性)：上記点推定では治療効果の臨床的な重要性を判断した。次には治療効果の区間推定値を吟味し、統計的に有意なのかどうかを判断する。通常区間推定には95%信頼区間を用いる。例えば、WOS 研究は一つの臨床試験である。もしも全く同じ臨床試験をもう一度行ったとしたら、全く同一の結果になるとは限らない。むしろ点推定は少し違ってしまおう。95%信頼区間とは、同じ臨床研究を100回行ったとしたらその100個の点推定のうち、極端な5個を除いた95個の点推定が入る範囲をいう。しかし、現実には1回しか研究を行っていない。そこで1回の臨床研究のデータから95%信頼区間を推定する必要がある。

最近の原著論文では95%信頼区間を報告しているものも多い。WOS 研究でも相対危険減少の95%信頼区間を報告している。しかし論文にいつも記載があるとは期待できない。WOS 研究でも絶対危険減少と NNT については点推定も区間推定も記載がない。この時に読者が電卓で区間推定を行うのは容易でない。

私たちは表計算ソフト EXCEL を使って、治療効果や診断テストの有効性の点推定と区間推定を簡単に行うことができるマクロ CLINICALC を用意している。希望の方は著者までご連絡をいただければ無料で提供している。ちなみに WOS 研究での絶対危険減少 2.2% と NNT45 の 95% 区間推定は、それぞれ 1.1~3.4%、29~92 である。

## 3) 研究結果の適用可能性

研究結果をあなたの患者に適用できるかどうかを決めるために、以下の点について考えてみる。

(1) あなたの患者は研究の対象となった被検者と似ているとってよいか?：あなたの患者は、年齢や性別、疾患の重症度、人種などの点で被検者とは異なっているかもしれない。研究結果をあなたの患者に当てはめてよいかどうかを判断するには、妥当性の検証のところで書き出した臨床試

験の資格基準を見ることが参考になる。ただし資格基準のすべてを満たしていなければならないと厳密に考えるのではなく、むしろ「この結果が役に立たないほど、私の患者はその試験の被検者と重要な点で大きく違っているだろうか？」と考えてみるのがよい。

(2) あなたの患者はその治療効果に満足するだろうか？：これに答えるためにはあなたの患者における治療効果の大きさを知る必要がある。臨床試験の結果は、治療効果の平均的な推定値を与えるものである。あなたはさらに、この結果から、あなたの患者で予想される治療効果を推定する必要がある。

ここでは治療効果のNNTがわかりやすい。NNTは何人の患者が治療を受けることで1例の悪い転帰を防止することができるかを示す数値であった。あなたの患者がもしも偽薬治療を受けたとしたら、WOS研究の偽薬群の患者に比べて虚血性心疾患を罹患する可能性がどの程度になるかを考えてみよう。

Kodamaら<sup>7)</sup>は日本人でのコホート研究を行い、血清コレステロール値が240mg/dl以上の女性での虚血性心疾患の罹患率は2.3/1,000人年と報告している。WOS研究では3,293人に偽薬を投与して平均4.9年追跡したので、その観察人年はおおよそ16,000(=3,293×4.9)であり、このうち248人が虚血性心疾患を発症したので罹患率(密度)は15.5/1,000人年である。したがって日本人女性でのリスクはWOS研究の偽薬群の、おおよそ2.3/15.5=0.15倍である。すると、あなたの患者でのNNTは報告されたNNTを0.15で割って求められる。WOS研究で報告されたNNTは45であったから、あなたの患者でのNNTは45/0.15=300である。さらにWOS研究でのpravastatinの投与量1日は40mgであり、これは日本での投与量の4倍である。したがって、あなたの患者でのNNTはさらに高いはずである。この数値にあなたの患者が満足するかどうか、患者さんとよく話し合ってみよう。

(3) あなたの患者は治療に伴う費用や危険の程度と頻度を受け入れるだろうか？：患者さんに治

療の話を進める際には、利益だけでなく不利益をも伝える必要がある。なかでも大切なのは、費用と治療に伴う危険であろう。Pravastatinの薬価は1錠(10mg)が187円である。またWOS研究では危険として心血管系疾患以外の原因による死亡、悪性腫瘍の発生、筋肉痛の症状、そして血液検査における異常所見の有無について、pravastatin群と偽薬群とで差がなかったことを報告している。これらを説明し治療の方針を決めてゆくことになる。

### 終わりに

治療に関するEBMというと新薬の臨床試験を思い浮かべてしまいがちだが、その病気の転帰を改善するために何ができるかと考えれば、手術、投薬以外にも、食事療法、運動療法も治療の介入に含まれる。善かれと思って行った治療が、害を及ぼしてしまうことのないように、本来すべての介入は検証されるべきである。今回、ワークシートに書き込む作業を通してWOS研究を批判的に吟味し、EBM実践の1例として紹介した。

1999年6月25日に行われた東京女子医科大学臨床判断学研究会で、WOS研究を行ったJames Shepherd先生のお話を聞く機会に恵まれた。これだけの大規模試験を行うためには、莫大なお金と時間がかかり、熟練した組織が必要という話であった。患者登録の点からも、経済的にも大規模な臨床試験を日本で実施するのは困難であろう。また、虚血性心疾患の発生率が日本の10倍近くあるスコットランドで行われたこの論文の結果を、すぐに我が国に当てはめるわけにはいかない。これは完成された論文ではあるが、すなわち、内的妥当性、重要性には優れているが、我々が日常診察する患者にはそのままは当てはまらない。

そのためにも日本人を対象とした臨床研究を、高脂血症の治療ばかりでなく、あらゆる疾患の診断、治療、予後について行う必要がある。そして将来的には、日本人のデータを元にした研究の集積から“我が国の”疾患に関するガイドラインが作られるべきであり、医師の勘や好みに左右されない統一見解が求められる。すでに出されているガイドラインについても、その勧告は妥当か検討

する必要があるかもしれない。情報公開, informed consents の普及に伴い, これから我々が取り組むべき課題である。こういった組織だった仕事は, 一機関で行うことは無理であろうが, 日々の診療において常に根拠を探る目を持ち続ける姿勢はこれからの臨床医にとって重要となる。

もちろん, 我々は論文の根拠に基づいた治療が, 自分の目の前の患者に最適とは限らないことも認識する必要がある。つまり, 経済的な問題, 社会的な問題, quality of life の問題等は画一的には評価しにくい, 主治医としては当然考慮するであろう。要するに, 十分に情報を提供した上で, 患者背景を考慮し最終的には患者の利益を優先した治療法を決定する。しかし, それまでの過程は経験や勘だけに頼るのではなく, “根拠に基づく医療” ということである。

#### 文 献

- 1) Oxman AD, Sackett DL, Guyatt GH: Users' guides to the medical literature. I How to get started. The Evidence-Based Medicine Working Group. *JAMA* **270**(17): 2093-2095, 1993
- 2) Guyatt GH, Sackett DL, Cook DJ et al: Users' guides to the medical literature, II How to use an article about therapy or prevention. A. Are the results of the study valid? *JAMA* **270**(21): 2598-2601, 1993
- 3) Guyatt GH, Sackett DL, Cook DJ et al: Users' guides to the medical literature, II How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? *JAMA* **271**(1): 59-63, 1994
- 4) 小島原典子, 岡本高宏, 香川 順: 公衆衛生学教室のクリニカル・クラークシップにおける Evidence-Based Medicine 導入の試み. 医教育: (in press), 2000
- 5) Shepherd J, Stuart MC, Ford I et al: Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Engl J Med* **333**(20): 1301-1307, 1995
- 6) 阿字川和雄, 後藤恵子, 加藤砂織ほか: Evidence-Based Medicine (EBM) 入門ガイド (2) 文献検索の方法論. *東女医大誌* **70**(1・2): 1-9, 2000
- 7) Kodama K, Sasaki H, Shimizu Y: Trend of coronary heart disease and its relationship to risk factors in a Japanese population: A 26-year follow-up, Hiroshima/Nagasaki Study. *Jpn Circ J* **54**: 414-421, 1990