

EBM は臨床医の新たな行動指針 ポイントは研究デザインと統計学的な妥当性

EBM (Evidence Based Medicine、証拠に基づく医療) の実践は今や、臨床医の新たな行動指針となったといっても過言ではない。EBM の推進により、医療の質の向上と標準化が図られ、地域や医師による診療内容のばらつきを極力減らすことが可能となる。また根拠のある臨床判断を患者に明示することによってインフォームド・コンセントも得やすく、医師と患者のよりよい信頼関係の構築にも役立つ。今後、EBM の手順で作成された診療ガイドラインに基づく診療がますます求められ、またそれがインターネット等で広く一般に公開されことにより、患者が標準的な診療内容の知識をもって、客観的で質の高い診療を期待するようになるのも時間の問題である。

かつて、「飲んだ、治った、効いた」の“3 た理論”によって医療行為が判定されていた時代もあった。医療の閉鎖性の中で患者も「そんなものか」と不安ながらも受け入れていた。しかし疫学的知識が蓄積され、バイアスやプラセボ効果などの問題が明らかとなり、こうした非科学的な判断は顧みられなくなった。

EBM は 1991 年、カナダのマクマスター大学で初めて使用され、概念が確立された。その定義は「診ている患者の臨床上の疑問に関して、医師が関連文献等を検索し、それらを批判的に吟味した上で患者への適用の妥当性を評価し、さらに患者の価値観や意向を考慮した上で臨床判断を下し、専門技能を活用して医療を行うこと」とされている（厚生労働省「医療技術評価推進検討会報告書」より）。

5 段階に分類されるエビデンスの質 EBM にもいろいろなレベルがあり、米国予防医療サービス特別研究班は証拠の質の程度によって 5 段階に分類している（表、Wolf SH, Sox HC. The expert panel on preventive services, continuing the work of the USPSTF. American Journal of Preventive Medicine 7:325-330, 1991）。

表 エビデンスの質の分類
(米国予防医療サービス特別研究班)

I	ランダム化比較試験
II-1	非ランダム化比較試験
II-2	コホート研究または症例対照研究
II-3	時系列研究、非対照実験研究
III	権威者の意見、記述疫学

これによると、「某先生の理論によれば」というのは最低レベルのエビデンスであり、患者を無作為化して複数の治療法や処理を割り付けて行うような研究デザイン (RCT) が最善のエビデンスだとしている。米国保健政策研究局 (AHCPR) は、複数の RCT を統合したメタ・アナリシスをさらに上位のエビデンス研究として位置づけている (1993 年)。とはいえ、質の悪い RCT よりよくデザインされたコホート研究やケース・コントロール研究の方が“証拠能力”が高いことも十分あり得る。従って、EBM を実践する際は、医学論文や学術発表を科学的・論理的に吟味し、EBM を見抜くことが必要となるのである。

臨床疫学に統計学が必須の理由現在では科学的な論文・発表は、臨床疫学としての研究デザイン (研究計画と) と統計学的に妥当な分析方法が必須である。海外の有力学術誌には臨床医学統計の専門家に

よる審査があり、統計学的に妥当な研究デザインによる論文でないかぎり、受け付けられない。医学分野に限らず、人間を対象とする調査研究では、科学的な方法として統計学を用いることは常識である。EBMの基礎となっている臨床疫学（疫学的な手法を用いて診断・検査・治療法の有効性と効率性を評価する学問）も、その基礎は統計学である。ではなぜ、科学的な研究に統計学を用いるのであろうか。

第一に、統計学は偶然による変動（ばらつき）を評価できるからである。人間集団のように個体差の大きな対象を扱う場合、比較するグループ間に対象の選び方による差（バイアス）が生じないように、まず研究デザインを正しく作成する必要がある。薬剤の比較試験などで用いられる無作為割付などはその一例である。疫学で用いるケースとコントロールの属性などのマッチングもそうである。そしてそのような状況であっても偶然性はあるので、それを評価しなければ正しい結論を得ることができない。その際、統計学のいろいろな分析方法が用いられるが、それらの方法のほとんどは、帰無仮説（証明したい仮説の逆の仮説、つまり結果として出てほしくない仮説）と呼ばれる仮定のもとで、その帰無仮説が偶然生じる確率を求めることで行われる。有意確率（p 値）とは、その偶然によって起きる仮説を推定しているのであり、その大きさを参考に意思決定が行われる。（例えば、「喫煙者は肺癌になりやすい」を証明したい仮説とすると、統計学的に分析した結果、 $p < 0.001$ という有意確率の数字が得られた場合、「喫煙者は肺癌になりにくい」という帰無仮説は1000分の1の確率であり、よって初めの仮説は相当正しいことが言えるわけで、肺癌予防のための禁煙指導は正しいことになる）。

統計学的視点のない研究は主観的で権威主義的臨床疫学に統計学が欠かせない第二の理由は、統計学が学問の特徴として、可能な限りあらゆる現象を数量化するからである。各種臨床検査値はもちろんのこと、患者の臨床症状、性格特性など“質”と考えられるものでも“量”として把握するよう工夫する。現象を量として扱うことは、その客観性を高めるために不可欠なのである。その結果、他の研究者がその結果について再度調査し、自分の研究と比較することを容易にすることができる。研究結果の再現性や比較可能性は、科学的な研究には欠かせないものであり、従って、統計学的な視点のない研究は、一般性（普遍性）を欠如した主観的、権威主義的、もしくは自己満足的なものになりがちである。これらの点から、科学的な根拠に基づく臨床疫学は、その方法論として必然的に統計学的アプローチが不可欠なのである。

本シリーズでは、学会発表や学術論文などの情報を科学的・論理的に吟味しEBMを見抜くためのポイントを、臨床統計学の視点から例題をまじえながらわかりやすく解説していきたい。

次号のテーマ（予定）は「“研究デザイン”なくして“臨床統計学”なし」。今述べたように、ある研究が科学的であるかどうかは、研究デザインが適切かどうか、その分析が統計学的に妥当なものかどうかにかかっている。何を知りたいかによって、その答えを与えてくれる研究デザインは異なるものであり、結果を導くための統計学的手法（結果の比較指標など）も違ってくる。研究デザインの持つ意味や研究目的別の研究デザインの種類、それらの統計学的な意味などについて触れたい。

なお、本ニューズレター「エビデンス」誌上ではサマリーを掲載し、詳しくはホームページ (http://www.novartis.co.jp/IT_on_Healthcare.htm) を読んでいただきたい。

高木 廣文氏 略歴

1950 年生まれ。74 年東京大学医学部保健学科卒業後、79 年同大大学院博士課程修了。

同年米国国立衛生院奨励研究員として、米国国立環境保健学研究所勤務。

81 年より聖路加看護大学講師（統計学）、83 年同大助教授。89 年文部省統計数理学研究所助教授、99 年 4 月新潟大学医療技術短期大学部看護学科教授、同年 10 月同大医学部教授保健学科転任。

主な著書は「ナースのための統計学；データのとり方生かし方」（医学書院）「多変量解析ハンドブック」（現代数学社）「健康科学とコンピュータ」（共立出版）など。