

書評

Jeremy Howick

The philosophy of evidence-based medicine (Wiley-Blackwell, 2011)

本書の著者 Jeremy Howick は、オックスフォード大学にある Centre for Evidence-Based Medicine に所属する哲学者である。EBM はその名の通り、証拠に基づく医療を意味し、特徴としては、臨床実践において治療法を決定する際に、医学的知識のみではなく統計的知識を重視することが挙げられる。この際、証拠の信頼性によって優劣を設けることも重要な特徴である。この特徴が明示的に表現されているのが、証拠のヒエラルキーという考え方である。EBM の基本的考え方は、今日では広く浸透していると言って差し支えないだろう。より信頼性の高い証拠に基づいて治療法を選択するという考え方は、それ自体としては何ら問題もなく、推奨されて当然であるようにも思われる。しかし、話はそう単純ではないことが、本書を読めば明らかとなるだろう。

それでは早速、内容の紹介に入る。本書は全 4 部から構成される。第 1 部 “Introduction” (1 章から 3 章) では、EBM に関する基本事項が確認され、著者の基本的な立場が示される。まず EBM の主張する証拠のヒエラルキーを確認した上で、歴史的変遷にも配慮しつつ、その主張の中心的要素として、(1) ランダム化試験 (randomized (controlled) trials) 及びそのシステマティック・レビューは観察的研究 (observational studies) よりも強い証拠を**一般に**提供する、(2) 比較臨床研究 (ランダム化試験と観察的研究を含む) はメカニズムに基づく推論よりも強い証拠を**一般に**提供する、(3) 比較臨床研究は専門家判断よりも強い証拠を**一般に**提供する、の 3 つが挙げられる。そして、これら 3 つが**一般に**支持され得るか、以降で検討されることとなる。その際の視点として著者は、臨床的決定において良い証拠であるためには、臨床的に有効な必要があると強調し、その条件として、(1) 効果が患者に適切であること、(2) 利益が損害を上回ること、(3) 実験結果が日常の実践に適用可能である (外的妥当性がある) こと、(4) その治療が最善の選択肢であること、を挙げる。

第 2 部 “Do randomization, double masking, and placebo controls rule out more confounding factors than their alternatives?” (4 章から 8 章) では、ランダム化試験が観察的研究よりも強い証拠を提供するという EBM 支持者の主張を検討している。著者は、EBM 支持者がランダム化試験の優位性を主張する際の根拠として、ランダム化、ダブル・マスキング、プラシーボ・コントロールが、ランダム化試験によってのみ可能で

あるという点を挙げ、これらについて実際に厳密な意味で実践されているか、また方法論的に優位であると言えるか、という視点から検討している。前者については、多くの例からその困難さが度々言及されているが、後者がここでの主たる議論となるため、以降では方法論的側面を中心に紹介していく。

4章は、第2部への導入部として位置付けられる。まず著者は、実験方法の優劣を検討する基準として、より強い証拠を提供する実験方法とは、より多くの交絡因子を排除できる方法であるとする。ここで交絡因子とは、(1) 結果に潜在的に影響を与え得る、(2) 実験群と対照群に不均等に分布している、(3) 実験による介入と無関係である、という性質を持つ因子のことである。このような因子が実験結果の信頼性に影響を及ぼすことは、誰もが認めるところであろう。そして、全ての交絡因子を枚挙し尽くすことは明らかに不可能であるため、EBM 支持者は、ランダム化試験によって観察的研究よりも相対的に多くの交絡因子を排除できると主張するのである。

5章では、ランダム化が検討される。確かに潜在的な交絡因子の全ては認識出来ないにせよ、現在のところ認識されている交絡因子を排除することは可能であるため、著者は、基本的にはランダム化の持つ相対的に高い証拠能力を認める。特に、観察的研究においては排除されない自己決定バイアス（患者の選好から生じるバイアス）と配分バイアス（実験群と対照群の不均等割付けから生じるバイアス）を排除できることは、重要であるとしている。しかしながら、EBM の証拠のヒエラルキーには、「ランダム化が観察的研究よりも良い証拠として位置付けられるのは、治療の効果が、交絡因子の複合効果に勝るときに限る」という修正を付け加える必要があるという。この条件を著者が設ける理由は、以前より指摘されているパラドックスを回避するためである。そのパラドックスとは、最も有効であるはずの治療法のいくつかは、証拠のヒエラルキーに照らして最良とされる証拠によって支持されないという事態のことである。その例として、窒息に対するハイムリッヒ法や、心停止に対する体外式除細動などが挙げられている。これらの治療法は、ランダム化を用いた証拠によっては支持されていないのである。また、外的妥当性については、ランダム化によって積極的に保証されはしないものの、観察的研究の方が優れているとも言えないとしている。

6章では、ダブル・マスキングが検討される。先のランダム化試験が配分段階における交絡因子を排除するのに対して、ダブル・マスキングは遂行段階における交絡因子に関わる。この段階で問題となるのは、患者と治療奉仕者が持つ、治療に関する知識を原因として生じる、様々な因子である。そして著者は、ここでもランダム化と同様に、観察的研究によっては排除されない因子を排除可能であるとして、基本的には

ダブル・マスキングは良い証拠を提供すると言う。排除される因子としては、信念効果（患者自身によって治療状況の報告がなされる際に、治療に関する知識のために生じる影響）、ドロップアウト（対照群に割り振られたことを知った患者の、実験からの離脱）、そして併用治療（対照群に割り振られたことを知った患者が、他の治療を併用して受けること）を挙げる。しかし、これもランダム化と同様に、明らかに有効な治療法が、最良の証拠によって支持されないという事態が生じる。例えば、アナフィラキシーショックに対するエピネフリンの投与が挙げられる。著者は、このような事例を排除するために、ダブル・マスキングが試験の質を高めるのは、「治療効果が劇的であり、患者や治療奉仕者の知識から生じる交絡因子の効果が、相対的に重要でないとき」に限るという注釈の設定を提案する。

このように、議論の構造はランダム化の場合とほぼ同様であり、この2つの結論から著者は、ランダム化試験は観察的研究に対して、前述の条件を踏まえた上であれば、方法論的価値があるという結論に達する。

7章と8章では、プラシーボ・コントロールが検討される。特に、7章では実践的側面、8章では方法論的側面を中心に検討される。まず著者は7章で、適切なプラシーボ・コントロールを、実験による治療に特有でない効果の全てを含み、かつそれ以上を含まないものとして定義する。これによって、実験群とプラシーボ群を比較した際に、実験による治療に特有の効果を、過不足なく得ることが可能となる。とはいえ、実践的限界のため、実際には適切なプラシーボ・コントロールを遂行することは困難である。そこで著者は、実験過程の詳細な記述と、コントロールしようと意図した特質を明記することを求める。これによって、少なくとも、プラシーボ・コントロールが適切に行われたか否かが判断できるためである。次に8章で、著者はまず PCTs (placebo-controlled trials) が、プラシーボ・コントロールを行わない ACTs (active controlled trials) に優越する根拠として、「プラシーボと非プラシーボとを区別する能力（これを著者は first assay sensitivity と呼ぶ）」と、効果の高い治療法と低い治療法とを区別する能力（これを著者は second assay sensitivity と呼ぶ）とが、PCTs にはあるが ACTs にはない」とする主張を検討し、最終的にこれを退けている。また、著者の定義する良い証拠の基準では、臨床実践において有用な証拠を提供することが重要となるが、標準的な患者、臨床家、政策立案者の関心は、新しい治療法がプラシーボよりも優れているかではなく、既に確立された治療法よりも優れているかである。そのため、PCTs が付加的な利益をもたらすことは稀となる。また、実践的側面を踏まえると、PCTs が実験による治療の絶対的効果 (absolute effect) を測れるとする主張は、

適切なブラシーボ・コントロールが遂行されるという、実際には稀にしか保証されない前提を基にしている点も指摘される。

第3部“Examining the paradox that traditional roles for mechanistic reasoning and expert judgment have been up-ended by EBM”(9章から11章)では、メカニズムに基づく推論と専門家判断の証拠としての信頼性が検討される。9章は、第2部のまとめと第3部への導入である。

5章から8章では、EBM支持者の主張する証拠のヒエラルキーにおいて、ランダム化試験を観察的研究に対して「一般に」優位に置くことの批判がなされたが、ここで扱われるメカニズムに基づく推論(10章)と専門家判断(11章)は、EBM支持者によってヒエラルキーの外に置かれるものである。これについて著者は、証拠能力を積極的には認められないという意味において、基本的にはEBM支持者の主張を認めている。本書ではそれぞれに一章ずつ割いて詳細に検討しているが、ここでは割愛する。結論だけ触れておくと、メカニズムに基づく推論のなかでも質の高いものに限っては、証拠としての能力を認められる場合もあるが、専門家判断には証拠としての能力は認められない。但し、専門家判断には証拠としての役割以外の、(1)患者の好みや事情を外的証拠と調和させること、(2)治療上の利益を最大化すること、(3)専門家の暗黙知を必要とする様々な技術を用いること、などを含む重要な役割が期待されるとしている。そして最後に、第4部“Conclusions”(12章)でこれまでの議論を簡潔にまとめ、本書を締め括っている。

以上、本書の内容を駆け足で紹介してきたが、最後に疑問に思う点を記しておく。本書の中心的主張は、第2部において議論された、証拠のヒエラルキーへの部分的な修正の提案である。こうした修正が必要とされる主な理由は、最も有効とさえ考えられている治療法が最良の証拠によって支持されないという事態であった。具体的には、ハインリッヒ法、体外式除細動、エピネフリンなどが挙げられた。著者による修正を受け入れることで、これらの事例を証拠のヒエラルキーと矛盾なく説明できることは、確かに魅力的である。しかし、実践的には明らかに有効性が問題とならない事例を説明することと引き換えに、ランダム化試験の証拠能力の優位性を保証するためには「治療効果が交絡因子の複合効果に勝る」という必ずしも明確とは言い難い条件を課すことになる。これによって証拠のヒエラルキーを修正することが、著者の強調する実践的意味において、どれほど有意義であるのかは疑問である。

最後に、本書ではこれまで紹介してきた内容が、豊富な事例や実験データと共に議論されている。そのため文献リストの充実は、これだけをとっても価値のあるもの

と言えるだろう。本書全体を通して展開される、実践的視点と方法論的視点からのバランスの良い議論と併せて、書名通り「EBM の哲学」なるものを概観するには有益なものとなっている。

(高尾克也, 京都大学大学院文学研究科修士課程 katsta07@me.com)

