

総説

画像診断のエビデンスを創るための研究デザイン

堀 雅敏

大阪大学大学院医学系研究科 放射線統合医学講座

Designing Diagnostic Imaging Research for Evidence-based Medicine

Masatoshi Hori, MD PhD

Department of Radiology, Osaka University Graduate School of Medicine

抄録

広い意味のEBMには「エビデンスを創る」、「伝える」、「使う」の三局面があるが、その第一の「エビデンスを創る」は臨床研究によって行われる。画像診断の臨床研究には様々な手法があり、その中にはエビデンス生成を主目的とすべきものとそうでないものとが含まれる。エビデンス生成を主目的とすべき画像診断研究には、診断法の評価・比較、メタアナリシス、システムティック・レビューが挙げられる。エビデンスを創る研究を行う為には、a) 良いテーマを選び、b) 適切なリサーチ・クエスチョンを想定して、c) 科学的に適切な手法で研究を行い、d) 適切に公表することが重要である。

Abstract

Clinical research makes evidences for medical decisions. There are various types of design in diagnostic imaging research. Among them, evaluation of accuracy of imaging tests, meta-analyses, and systematic reviews should be used for making evidences. It is important for researchers to (a) pick a good theme, (b) make appropriate research questions, (c) use scientific methods, and (d) publish research appropriately.

Key words EBM, evidence-based medicine, research design, diagnostic imaging, radiology

(1) はじめに

Evidence-based Medicine (EBM、根拠に基づく医療)は1990年代より急速に広まった概念・手法であり、元来は個々の患者を対象とした個々の臨床状況での意志決定において、最新で最善の根拠をその他の個別要因と共に利用するというものとして規定

されている。このように、エビデンスを「使う」状況での考え方が本来のEBMであるが、一方で広い意味のEBMには「エビデンスを創る」、「伝える」、「使う」の三局面があるという¹⁾。現在、様々な疾患を対象としてエビデンスに基づく診療ガイドラインが精力的に作成されている。画像診断の分野でも日本医学

刷刷請求先：〒565-0871 大阪府吹田市山田丘2-2 D1

大阪大学大学院医学系研究科 放射線統合医学講座 堀 雅敏

TEL：06-6879-3434 FAX：06-6879-3439

放射線学会および日本放射線科専門医会・医会合同ガイドライン委員会によって「エビデンスに基づく画像診断ガイドライン」が編纂されつつあり、完成した部分がWebページ上で公開されている (<http://www.jcr.or.jp/guideline/guideline.html>)。こうした診療ガイドライン作成はエビデンスの収集・評価・提供を行うプロセスであり、前述のEBM三局面のうち「エビデンスを伝える」に該当するという¹⁾。それでは、EBM三局面の最初にある「エビデンスを創る」とは、どういうプロセスであろうか？ これは取りも直さず臨床研究を行うことである。従って、「エビデンスを創る」というようなことを大げさに考える必要は無く、今まで通りに臨床研究を行い、学会で発表し、論文としてその内容を公表してゆけばよいというふうにも考えられる。しかし、それでよいのだろうか？ 少し考えればすぐに分かることであるが、あらゆる臨床研究が同等に「エビデンスを創る」わけでは無い。我が身を振り返ってみると、臨床研究を行う際に、その研究内容を診療エビデンス生成の観点から考えたことはあまりなかった。そんなこともよく考えずに臨床研究をしていたのか…とあきれ先生方もいらっしやることと思うが、私と同様に、十分に考えてこなかった先生方も大勢いらっしやるのではなからうか。本稿では、画像診断分野における臨床研究について、「エビデンスを創る」観点から述べる。

(2) エビデンスはどのようにして創られるのか？

エビデンスを創る研究デザインを考える前に、エビデンスがどのようにして創られるのかを考えてみたい。EBMが一般的になる以前、臨床的な疑問が生じたときに行われたことは、教科書を調べる、あるいは適当な論文が無いか検索してそれを読むことであったと思われる。現在でもしばしば我々はこの方法で臨床的疑問に対処する。一般に、教科書はその道の研究者・臨床家が自らの専門的知見に基づいて記載しているものであり、文献の裏付けの存在する場合も存在しない場合もある。エビデンスについて明確に考慮して執筆されてきたものは少なく、今日の科学的なエビデンス基準から考えれば一般にそのエビデンス・レベルは低いと考えられる。それでは、論文はどうであろうか？ 論文は科学的な体裁を整えて出版されている。すなわち、目的、方法、結果、考察が明確に分けて記載されており、そこか

ら出てきた結論は科学的真実であるようにみえる。しかし、臨床研究の論文では、目的はよく似ているのに結論はかなり異なっている場合も出てくる。そうした場合は、どの論文の結論を科学的に妥当だと考えるべきなのか問題になってくる。簡単な対処法としては、論文の掲載誌のレベルで判断する場合もあるだろう。例えば、雑誌Radiologyに掲載されている画像診断の論文の結論が、他の雑誌に掲載されている論文の結論と異なっている場合は、Radiology論文の結果を無条件に受け入れるというようなやり方である。Radiology以外の雑誌を比較する場合には、雑誌のインパクト・ファクターで比較すればよい。簡単な手法であり、いつでも迷わずに短時間で行うことができる。しかし、これは科学的な態度とは言えず、EBMの実践とは言い難い。EBMでいうエビデンスは、これまでに公表された論文個々の結果・結論を無批判に利用して得られるものではない。公表されている論文の結果・結論が、自らが今から行おうとしている医療に適用可能かどうかを批判的に吟味することがEBMでは求められる。論文が取り扱っている対象者は、自分の患者集団とそれほど変わりがないかどうか？ 論文の研究手法は適切といえるのか？ この研究はどの程度確からしいと言えるのか？ これらを吟味し、吟味した結果が適切であればエビデンスとして診療計画に用いることになる。科学的研究といえども、結果には様々なバイアスが入ってくる。他の分野の科学論文に比べて、臨床研究ではこのバイアスが大きくなりがちであるが、それを適切に評価してエビデンスとする。日常の診療の合間に、そのような吟味を常に行うのは困難であるから、それを助ける目的で各種の「EBMを用いた診療ガイドライン」が作成されている。個々の医師が独力でEBMを実践する場合にせよ、EBMを用いた診療ガイドライン作成を行う場合にせよ、エビデンスを考えるとときには批判的吟味が行われる。文献の批判的吟味とは、研究の科学的妥当性の評価と、結果の臨床的重要性を吟味し、その文献が実際の問題解決に役立つかどうかを評価するものと言われている¹⁾。このように考えると、「エビデンスを創る研究」の必要条件のひとつが見えてくる。それは、上記のような批判的吟味に耐える研究を実施し、公表することと言えよう。ここでは、研究を実施することと結果を公表することは、別の事柄であることを強調したい。これらは当たり前のことのよう

に思われるが、実は難しいことも多く、画像診断研究の分野でエビデンス・レベルの高い研究が少ないと言われるもとなっていると考える。

「エビデンスを創る研究」を考えると、もう一つ忘れてはならないことがある。EBMは臨床に直結した概念であり、エビデンスは臨床現場における患者のマネジメントに際して生じる臨床上の疑問(クリニカル・クエスチョン：CQ)に答えるものでなければならないということである。数あるクリニカル・クエスチョンのうち、倫理的、経済的、臨床的に様々な条件を満たしたものはリサーチ・クエスチョン(RQ)となり、研究が行われる²⁾。ここでは、医学論文が必ずしも臨床上の疑問に答えるものとは限らないことに注目する必要がある。顕著な例は基礎医学研究である。遺伝子レベル・細胞レベルの基礎医学研究は盛んに行われており、こうした研究のうちレベルの高いものはインパクト・ファクターの極めて高い雑誌に掲載され、一般に価値の高い研究とみなされている。こうした基礎医学的研究の結果は科学的エビデンスのひとつには違いないが、EBMでいうところのエビデンスとは見なされない。医療の質に直接貢献するのは人間を対象とした臨床試験や疫学研究の知見なのである^{1,3)}。それでは画像診断を含む臨床医学の論文であれば何であってもEBMのエビデンスになり得るのかというと、そうでもない。これをみるために、リサーチ・クエスチョンの実例をみてみたい。次に挙げるのは、日本医学放射線学会および日本放射線科専門医会・医会合同ガイドライン委員会(JRS/JCR合同ガイドライン委員会)編エビデンスに基づく画像診断ガイドライン-2007(<http://www.jcr.or.jp/guideline/guideline.html>)の中から、肝海綿状血管腫に対する画像診断ガイドラインのリサーチ・クエスチョンを抜き出したものである。EBMとガイドラインは必ずしも結びつくものでないが、ガイドラインに取り上げられたリサーチ・クエスチョンを概観することはエビデンスを創るための研究を考える上で有用であろう。

肝海綿状血管腫の画像診断ガイドラインのリサーチ・クエスチョン(RQ)

超音波(US)

1. Bモード超音波は血管腫の検出に有用か
2. Bモード超音波で境界明瞭な高エコー結節は血管腫に特異的か

3. Bモード超音波検査中のエコー輝度の経時的変化は特異的所見か
4. Bモード超音波で血管腫の典型的所見が得られた場合、他の検査は必要か
5. 造影超音波は診断の決め手となるか

CT

1. CTは肝血管腫の検出に有用か
2. CTにおける典型的所見はあるか
3. Dynamic CTは肝血管腫の確定診断に有用か

MRI

1. MRIは血管腫の検出に有用か
2. MRIの撮像法は何かが良いか
3. 血管腫の診断に、ガドリニウム造影剤によるdynamic MRIは有用か
4. SPIO造影MRIは、他の一般的な画像検査で非典型的所見を呈する血管腫を、悪性腫瘍と鑑別する際の決め手となるか
5. 血管腫として非典型的な所見を示す腫瘍の鑑別診断には、MRIが有用か

確定診断

1. 肝血管腫の確定診断に最も信頼性の高い検査法は何か
2. 肝血管腫の診断に腫瘍生検は必要か

鑑別診断

1. USでの高エコーは血管腫に特異的か
2. MRI T2強調像での高信号は血管腫に特異的か
3. 造影CT/MRIでの早期辺縁結節状濃染は肝血管腫に特異的か
4. 造影CT/MRIでの後期全体濃染や中心部への濃染の広がり肝血管腫に特異的か

臨床

1. 肝血管腫の経過観察は必要か
2. 治療が必要となる血管腫はあるか

以上のように、リサーチ・クエスチョンは、臨床に直結した疑問である。ある検査法は診断に有用か？ある検査法により確定診断が可能か？ある所見は疾患に特異的か？エビデンスを創る研究は、これらの疑問に答える必要がある。これらに答えられない研究は、エビデンスを創るのにはあまり役立たないと言っていることができる。例えば、MRIの研究で2つのシーケンスを比較評価する場合を考える。このとき病変のコントラストについて調べたところ、シーケンスAの方がシーケンスBよりも有意に優れていた

という結果が得られ、対象患者、撮像方法、統計的手法について非常によく練られて研究が実行されたとする。これにより得られた結論は科学的見地から考えて、間違いのない事実であろう。しかし、これをもってシーケンスAの方がシーケンスBよりも優れた検査法であるかと問われると、何とも答えられない。コントラストが有意に優れていることは、診断能が有意に優れていることに直結しないからである。エビデンスを創ろうと思えば、画像のコントラストではなく、診断能を評価しなければならない。患者が検査を受けるときに重視しなければならないデータは、画像のコントラストではなくて、実際の診断能の違いでなければならない。これを高血圧の治療に関する疫学の例で考えると、降圧薬服用によって高血圧合併症やそれによる死亡の可能性を下げるためになるかどうか重要であり、血圧を下げるのが直接重要なのではないことと同様と考えられる。

(3) 画像診断臨床研究の種類とエビデンス

画像診断臨床研究では、次のような研究が行われている：

- (a) 画像診断法の評価・比較
- (b) メタアナリシス
- (c) システマティック・レビュー
- (d) 画像と病理の対比
- (e) 症例集積
- (f) 症例報告
- (g) 新技術報告
- (h) その他

このように多くの臨床研究手法が存在するが、これらの中には「エビデンスを創る」ことに直接つながりやすい研究手法とそうでない研究手法がある。上記のうち、(a)～(c)は「エビデンスを創る」研究に直結しやすいし、またそのように研究を進めるべきものである。一方、(d)～(g)はエビデンスにはなりにくい研究である。ただし、エビデンスを創ることに役立つ研究だからと言って価値がないわけではないことにも注意しておきたい。画像と病理の対比は、画像で見えている所見が、どのような病理に対応しているのかを明らかにする研究であり、多くの画像診断医が興味を持ち、かつ画像診断の進歩に役立ってきた。症例集積や症例報告もめずらしい画像所見・特徴的画像所見を呈するものであれば、しばしば実際の臨床にとっても役立つことを

体験する。基礎医学的研究がEBMでいうところのエビデンスにつながらないと同様、「重要な」あるいは「役に立つ」臨床研究の全てがエビデンスにつながるわけではない。従って、エビデンスにならない研究をするべきでないという主張は明らかな間違いであろうと私は考える。とはいえ、本稿は「エビデンスを創るための研究」がテーマであるので、以下では(a)～(c)について述べる。特に「(a) 画像診断法の評価・比較」について詳述する。

(4) 各種の画像診断研究手法

…(a) 画像診断法の評価・比較

画像診断法の評価・比較としては、診断精度(検出、鑑別)、測定量の精度(臓器体積、血流速度、組織灌流量など)、診断の再現性(読影者内・読影者間あるいは施設内・施設間)、診断法の安全性やコスト、検査が治療計画策定に及ぼす影響、予後予測、検査が受診者の予後を改善するか否かなどについての評価・比較が行われる⁴⁾。エビデンスを創る研究の中では数が最も多く、多様であり、またエビデンスを創るのに不可欠の研究である。

診断精度(検出、鑑別)に関する研究

その画像検査が正しい診断結果をもたらす程度を評価する研究である。この研究では、画像検査の結果を正しい答え(ゴールド・スタンダード)と比較することが行われる。研究デザインで使用される統計手法として、感度、特異度、陽性・陰性予測力、尤度比、ROC解析などがある。

測定量の精度に関する研究

画像診断検査で、臓器体積、血流速度、組織灌流量など連続値をもつ物理量が測定されることがある。こうして得られた測定量の精度を評価する研究であり、他の手法で求められた値と比較する。この比較を行う際に、しばしば二法による測定値の相関係数を調べるが行われるが、これは適切でないとして、避けるべきである。適切でない理由として、相関係数ははずれ値の影響を強く受けることなどが指摘されている。異なる手法による測定値を比較する方法として、Bland-Altmanプロットが推奨されている⁵⁾。これは、二法によって得られた値の平均値をx軸に、それらの差をy軸にプロットする方法である。

診断の再現性に関する研究

読影者内・読影者間あるいは施設内・施設間での診断の再現性を評価する研究である。ある読影者が時間をおいて同じ画像を二度読影するとき、二度の読影結果がどの程度一致するかということが観察者内再現性の問題である。また、同じ画像を別の読影者が読影するときに読影結果がどの程度一致するかということが観察者間再現性の問題である。これらの一致が悪い場合には有用な画像診断法とは言えないことになる。これらの評価には κ 統計が用いられる。

診断法の安全性やコストに関する研究

造影剤の副作用頻度を評価する、あるいは検査の平均コストを比較するなどの研究である。平均、標準偏差、中央値、範囲などを算出する。

検査が治療計画策定に及ぼす影響に関する研究

検査を行うことで治療計画策定に及ぼすか否かを検討する。研究デザインとしては、検査前に治療計画を記載しておき、検査後に策定される治療計画と比較する。治療計画の変更につながった検査結果の割合を示すことになる。

予後予測に関する研究

検査によって得られた情報が、予後の予測に有用であった度合いを評価する

検査が受診者の予後を改善するか否かに関する研究

ある検査を受けた人のアウトカム(生存期間など)が、検査を受けなかった人に比べて良いか否かを評価する研究である。ランダム化臨床試験を行うのが望ましいが、実際には困難なことが多く、ケース・コントロール研究やコホート研究が行われる⁴⁾。

画像診断法の評価・比較に関する研究ではエビデンスを生成に役立つことを目指して、研究を計画し、実施し、公表することが望まれる。それには、1) 臨床の疑問に直接答えられるようなテーマを選び、2) 適切な方法を用い、3) 研究が批判的に吟味されることを想定して公表すること、が必要である。第一番目の「適切なテーマ選び」については、豊かな臨床経験と新しい技術に対する適切な理解が助けになるだろう。テーマ選びにおいては、「こんな技術が出てきたから何か研究してみよう」という方向でなく、

「こんな臨床的疑問は、この技術でどのように解決できるのか」といった方向から考えることが重要であろう。新たな技術が得られてから、臨床的疑問を考えるのでは、良いテーマを設定するのは困難であろうと思う。良いテーマを考え付くには、日常臨床を通じ、学会に参加しながら多くの臨床的疑問を常に意識していることが重要であろう。第二番目の「適切な方法を用いる」ことは、研究の一番難しいところであろう。研究論文の価値は、その論文が扱った特定の対象者集団で得られた結果を通して、他の一般的な診療において有益な情報をもたらすか否かによって決まる。エビデンス・レベルの高い研究を実施するためには、方法をよく検討して、研究のバイアス(偏り)を最小にし、正しい統計手法を用いて結果を解析することが重要である。画像診断研究にエビデンス・レベルの高い研究が少ないと言われるのは、この部分が十分に行われていないことが多いためであろう。困難な点として、第一に選択バイアス(対象者が恣意的に選択されるなどして生じるバイアス)が最小化されていない研究がよくみられること、第二にサンプル・サイズの少ない研究が多いこと、第三にしばしば統計手法が適切でないこと、などがよく問題になる事項である。例えば、二つの画像診断法A,Bの診断能比較を行う場合に、次のような研究が行われることがある:「2005年～2008年の間に、ある病院で画像診断法Aと画像診断法Bの検査を受け、さらにゴールド・スタンダードを得るための検査も行われたXX名を対象として、画像診断法AとBの優劣を比較した」。このような研究プロトコルでは、これより後の研究段階を非常に苦勞して、注意深く実施したところで、エビデンス・レベルの高い結論を得ることはできない。何故だろうか? それは、この研究の対象患者の選び方に大きな選択バイアスが生じている可能性を排除できないからである。まず、これらの検査を受けた人たちがどういう理由で検査を受けたのかが分からない。さらに、画像診断法Aだけを受けてBを受けなかった人の数はどれだけでその理由はどういうことだったのかが分からない。もし、ほとんどの人は画像診断法Aを最初に受けて、その中から精査を必要とされた人だけが画像診断法Bを受けたのだとすると、このような研究が、画像診断法AとBを公平に比較して、適切な結果をもたらすとは思われないのである。また、ゴールド・スタンダードを得るための検査はど

のような人に対して行われたのかが分からないのも問題である。画像診断法A,Bともに結果が陰性だった人に対してはゴールド・スタンダードとなるべき検査をしなかったのだとしたら、検査陰性の人の中に実際にはどれだけの見逃しがあったのか見当も付かないことになる。研究方法の最初の2~3行を見るだけで、この研究のエビデンス・レベルが高くないことが分かってしまうのである。実際にはこのような研究手法は画像診断の論文にしばしば見られる。エビデンス・レベルの高い研究を行うためには、上記のような様々な事項が曖昧であってはならない。

第二のサンプル・サイズの問題は画像診断の臨床研究にとっては悩ましい場合が多い。一般的には、研究計画の段階で必要なサンプル・サイズを見積もり、その数の臨床例を集めてデータ解析を行うべきである。しかし、このようにして見積もられたサンプル・サイズが大きくて、一つの病院の研究グループでまかないきれない場合も考えられる。また、新しい試みの場合はサンプル・サイズの見積もり自体が困難な場合もある。このような場合の対処法であるが、本当にエビデンス・レベルの高い研究を目指すのであれば、複数の病院での共同研究に持ち込むことが考えられる。但し、このようなことはいつでも可能とは言えない。サンプル・サイズが小さくても出版可能であれば、公表するのが良いだろうと私は考える。後述のメタアナリシスでは、個々の研究のサンプル・サイズが小さくても、多数の研究を評価することで、よりレベルの高いエビデンスを得ることが可能であり、例えばサンプル・サイズの小さな研究であっても、そうしたメタアナリシスを通じて、レベルの高いエビデンスを創ることに貢献し得るからである。第三の問題である統計手法の誤りについては、改めて言うべきことでもないであろう。正しい統計手法が用いられていない研究の結果は非常に疑わしいものとなることは論をまたない。

(5) 各種の画像診断研究手法

…(b)メタアナリシスおよびシステマティック・レビュー

メタアナリシスとは、同じ研究テーマについて実施された過去の多くの研究をひとまとめにして統計的手法を用いて再解析する研究手法をいう。この手法は、ある研究テーマに関する研究結果について、論文間に食い違いがある時や、個々の研究のサンプル・サイズが小さすぎて単独では結論づけられ

ないような時に有用であるとされる⁴⁾。また、システマティック・レビューは、「明確に作成された疑問に対する文献レビューで、系統的・明示的方法により、適切な研究論文を同定・選択、批判的に吟味し、そしてレビューに含まれた各研究が示すデータを収集し分析するものである。統計的手法(メタアナリシス)は、結果の分析や集約に用いられる場合も、用いられない場合もある(コクランレビューマニュアル)」と定められている¹⁾。一般的には、メタアナリシスやシステマティック・レビューは個別の研究よりもエビデンス・レベルが高いと考えられていることが多いが、これに反対する意見も存在する¹⁾。

(6) 研究の公表

苦勞して行った研究であっても、公表しなければエビデンス生成に全く寄与しないから、研究を公表することはエビデンスを創る上で極めて重要である。忘れがちであるが、単に研究を公表するのみでなく、適切に公表することが重要である。公表された研究がエビデンスになるには、その研究が批判的に吟味されなければならない。論文に記載された結果・結論がそのままエビデンスとして採用されるのではないのである。従って、研究を公表する論文には、読者が研究方法について吟味できるだけの十分な情報が記載されていなければならない。それではどれだけ書き込めばよいのだろうか？ この観点から、診断能評価の研究論文についての要求を25項目からなるチェックリストの形でまとめたものが the Standard for Reporting of Diagnostic Accuracy (STARD) initiativeとして提唱されている。これは画像診断の論文のみならず、様々な診断法に関する論文を対象として考案され、2003年にRadiologyを含む7つの医学雑誌上で同時に公開された⁶⁾。その後もAmerican Journal of Roentgenology (AJR) など幾つかの雑誌に掲載されるとともにWebページでもその内容が参照可能となっている (<http://www.stard-statement.org/>)。RadiologyやAJRなど多くの雑誌の執筆ガイドラインには、「診断精度を取り扱う研究については、STARDチェックリストを参照すること」が求められている。このチェックリストは、研究に含まれるバイアスの可能性を論文の読者が評価できるようにし、さらにその研究によって得られた知見を自らの患者に適用することが可能か否かを見積もる為に必要な事柄が

選び出されている。title/abstract/keywords、introduction、methods、results、discussionのセクション毎に書き込むべき内容が一覧表になっており、論文を記載する上で大いに参考になる。また、アクセプトされやすい論文を書く上でもこのリストは役に立つだろう。

(7) ケース・スタディ

1. 大腸癌肝転移の検出診断能：CTとMRIの比較

大腸癌患者に肝転移があるか否かについて、CTとMRIのどちらが正確に診断できるかを調べたいとする。どのような研究プロトコルが望ましいだろうか。「2005年～2008年の間に、ある病院でCTとMRIの両検査を受けた大腸癌患者を診療記録から検索して対象とした」というやり方では、エビデンス・レベルの高い研究にならないことは前述の通りである。この病院では大腸癌患者全員がCT検査を受けたが、その結果で異常が疑われた場合にMRI検査が行われる傾向があり、また異常所見が無かった場合はMRI検査が行われない傾向があったのかもしれない。そのような場合はバイアスがかかる為である。このようなバイアスが研究結果にどのような影響をもたらすのか考えてみる。表1はCT検査の結果（肝転移あり/なし）と真の結果（肝転移あり/なし）を示す分割表である。この場合、肝転移検出におけるCT検査の感度は $a/(a+c)$ 、特異度は $d/(b+d)$ である。CTとMRIの両方の検査が行われた患者のみを対象として感度と特異度を評価するときにはどのようなになるだろうか？ CTで肝転移なしとされた場合の方が、CTで肝転移ありとされた場合に比べてMRIが施行される割合が低いとすれば、CTで肝転移なしと診断された患者はMRIが行われず、研究対象からはずれる傾向が強くなることになる。そうすると、表1のうちaよりもcの減少が大きく、bよりもdの減少が大きくなることになる。この時、感度 $a/(a+c)$ は上昇し、特異度 $d/(b+d)$ は減少する。このような研究デザインでは、CTの感度は良く、特異度は悪くなってしまふのである。こうしたバイアスを避けるためには、研究したい検査法による検査結果が研究対象者の選択に影響しないように注意しなければならない。本当に診断能比較を行いたいのであれば、この病院の大腸癌患者全例または大腸癌で手術する予定のある全例を対象として、CTとMRIの両方の検査を施行して比較す

るべきであろう。一方で、このようにすると患者の負担が増加することになる。それでも、臨床研究として実施する価値があると考えられる場合に限り、研究プロトコルを倫理審査にかけて承認を受け、対象となる患者へ説明し、対象者の同意を得て、研究を実施する。

表1. CT検査の結果と肝転移の有無

	真に肝転移あり	真に肝転移なし
CTで肝転移あり	a	b
CTで肝転移なし	c	d

Note. 表の中のa,b,c,dは症例数

2. CTコロノグラフィ(大腸仮想内視鏡)による大腸ポリープの検出診断能

CTコロノグラフィによる大腸ポリープの検出診断能を評価する研究を考える。検出診断能の評価であるから、ゴールド・スタンダードとの比較が必要である。ゴールド・スタンダードとして、大腸内視鏡検査が使用されることが多い。前のケース・スタディの時と同様に、CTコロノグラフィの検査結果によって大腸内視鏡検査の実施される割合が変わるようであれば大きなバイアスが加わることになり、エビデンス・レベルの高い研究にはならないので注意が必要である。

研究を行おうとするテーマによっては、研究手法や論文の作成について特別に注意が喚起されていることがあるが、CTコロノグラフィはまさにそのようなテーマである。HalliganらはCTコロノグラフィによる大腸ポリープと大腸癌検出についてのシステムティック・レビューとメタアナリシスを行い、これまでに公表されてきた論文の記載が十分でないことが多かったと報告し、最低限論文に書くべき事柄を示した⁷⁾。CTコロノグラフィでは、検査手法、コンピュータ表示法、読影方法(2Dか3Dか)、読影者の経験レベルなど、検査法を評価するために重要な情報が多い。このように、ある診断方法の研究に対して特別な提言が行われている場合には、その提言をふまえて研究計画・論文の構成を立案しておかなければならない。

3. 画像による健診の有用性を評価する

健診の有用性を評価したいとき、どのような研究を行えばよいだろうか？ ある健診で見つかった胃癌では早期癌の割合が約7割、自覚症状などを契機として病院を受診し胃癌が見つかった場合は早期癌の割合が約5割であったとする。健診では早期癌が見つかる割合が高く、また早期癌の5年生存率は95%以上であって予後がよい。従って健診を受けるべき…と推奨できるのだろうか？ 川村はこのような考えは間違いであり、エビデンスにならないことを示している⁸⁾。この考え方が間違いである理由として、リードタイム・バイアスが挙げられている。癌で症状が出てくるまでにはある程度の期間を要する。従って、たとえ死亡する時期が健診受診の有無によらず変わらないとしても、健診で癌を発見した時点から死亡までの期間は、症状が出てから癌を発見し

た場合に比べて長くなるのがあたりまえである。すなわち、受診者の真の生存期間を全く延ばさなくても、癌を早く見つければ見かけ上は生存期間が長くなってしまふのである。このように、発見された癌の中での早期癌の割合を比較することは適切でない。健診受診者と非受診者の間にアウトカム（死亡やQOLなど）の違いがあるか否かを評価しなければならない。

(8) まとめ

画像診断の臨床研究には様々な手法があり、その中にはエビデンス生成を主目的とすべきものとそうでないものがある。エビデンスを創る研究を行う為には、1) 良いテーマを選び、2) 適切なりサーチ・クエスチョンを想定して、3) 科学的に適切な手法で研究を行い、4) 適切に公表することが重要である。

参考文献

1. 中山建夫：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。
2. 福井次矢，吉田雅博，山口直人編：Minds臨床ガイドライン作成の手引き2007。医学書院，東京，2007。
3. Muir Gray JA. Evidence-Based Healthcare : How to Make Health Policy and Management Decisions. Churchill Livingstone (London) 2002.
4. 木原雅子，木原正博訳：Hulley SB，他著：医学的研究のデザイン 研究の質を高める疫学のアプローチ 第3版。メディカル・サイエンス・インターナショナル，東京，2009。
5. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1 (8476): 307-310, 1986.
6. Bossuyt PM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy : The STARD Initiative. Radiology 226: 24-28, 2003.
7. Halligan S, et al. CT colonography in the detection of colorectal polyps and cancer: systematic review, meta-analysis, and proposed minimum data set for study level reporting. Radiology. 237: 893-904, 2005.
8. 川村孝：エビデンスをつくる 陥りやすい臨床研究のピットフォール。医学書院，東京，2003。

ダウンロードされた論文は私的利用のみが許諾されています。公衆への再配布については下記をご覧ください。

複写をご希望の方へ

断層映像研究会は、本誌掲載著作物の複写に関する権利を一般社団法人学術著作権協会に委託しております。

本誌に掲載された著作物の複写をご希望の方は、(社)学術著作権協会より許諾を受けて下さい。但し、企業等法人による社内利用目的の複写については、当該企業等法人が社団法人日本複写権センター（(社)学術著作権協会が社内利用目的複写に関する権利を再委託している団体）と包括複写許諾契約を締結している場合にあっては、その必要はございません（社外頒布目的の複写については、許諾が必要です）。

権利委託先 一般社団法人学術著作権協会

〒107-0052 東京都港区赤坂 9-6-41 乃木坂ビル 3F FAX：03-3475-5619 E-mail：info@jaacc.jp

複写以外の許諾（著作物の引用、転載、翻訳等）に関しては、(社)学術著作権協会に委託致しておりません。

直接、断層映像研究会へお問い合わせください

Reprographic Reproduction outside Japan

One of the following procedures is required to copy this work.

1. If you apply for license for copying in a country or region in which JAACC has concluded a bilateral agreement with an RRO (Reproduction Rights Organisation), please apply for the license to the RRO.

Please visit the following URL for the countries and regions in which JAACC has concluded bilateral agreements.

<http://www.jaacc.org/>

2. If you apply for license for copying in a country or region in which JAACC has no bilateral agreement, please apply for the license to JAACC.

For the license for citation, reprint, and/or translation, etc., please contact the right holder directly.

JAACC (Japan Academic Association for Copyright Clearance) is an official member RRO of the IFRRO (International Federation of Reproduction Rights Organisations).

Japan Academic Association for Copyright Clearance (JAACC)

Address 9-6-41 Akasaka, Minato-ku, Tokyo 107-0052 Japan

E-mail info@jaacc.jp Fax: +81-33475-5619