第2回 疫学用語の確認と論文の読み方（疫学各論1）

小松 裕和^{a) b)} 鈴木 越治^{b)} 土居 弘幸^{b)}

要 旨

論文の読み方（批判的吟味 critical appraisal）は研究を行う上で基本的な技術であり、最近ではEBMの実践にあたり臨床医にも求められる能力になっている。論文を読むことを苦手とする臨床医は非常に多いが、筆者らは疫学の基本的知識と論文の書き方について勉強することが、論文を読む技術を習得する一番の近道ではないかと考えている。疫学の基本的知識としては、曝露（治療）のアウトカムへの影響を定量的に推定するために必要な疫学指標と効果指標についての理解、バイアスを考えるのに必要な研究デザインについての理解が非常に重要である。一方、論文の読み方に関しては、論文の段落構造を理解し、段落ごとに読んでいく「パラグラフリーディング」が行えるようになれば、論文を読む効率は飛躍的に向上する。そして、論文の書き方に関する各種ガイドラインを参考にしながら、論文を読んでいくことは一番のトレーニングである。

（日救急医学会誌 2009；20：338-44）

キーワード：論文の読み方、疫学指標、効果指標、研究デザイン、パラグラフリーディング

臨床研究を実施するに際しても、研究データを論文にまとめるに際しても、論文の読み方（批判的吟味critical appraisal）は最も基本となる技術である。そして、この技術は臨床研究を行わない多くの臨床医にもEBMを実践するために求められるものであり、問診や身体診察と同様に全ての医師に必要な基本技術になってきている。しかし、論文の批判的吟味はEBMの実践を目指す臨床医が増えてきた現在においても、未だに大きな壁となっている。巷にEBM関連の書籍や雑誌があふれる中、なぜ論文の批判的吟味の習得は進まないのだろうか。これは基本となる疫学知識の欠如とともに、論文の書き方について知る機会が少ないために起きている現象ではないかと、筆者らは考えている。

本稿では、臨床研究の実施の基礎となる論文の読み方（批判的吟味）を習得するために、疫学用語の基礎知識¹⁻³⁾についての整理をまず行い、その後に論文の書き方に関するガイドライン⁴⁻⁶⁾を参考に論文の読み方（批判的吟味）に関する解説を行う。

✦ 論文を読むために基本となる疫学知識 ✦

1. 臨床研究は仮説を検証するために行う

何よりもまず、多くの臨床研究は「仮説の検証」を行うために実施されているということを理解して欲しい。確かに、グループインタビューを通じて主要な概念を抽出する質的研究、多数の測定データから

a) JA 長野厚生連佐久総合病院地域ケア科

b) 岡山大学大学院医歯薬学総合研究科疫学・衛生学分野

相関関係や関連性を検証して仮説を生み出す研究、質問紙表の妥当性を検証する研究などもあるが、基本となるのは「仮説の検証」である。

若干補足すると、臨床研究のプロセスにおいて、現場での経験・観察や先行研究の知見から研究仮説（新しい考え）をまず立てる、この際に用いる考え方を帰納法と呼ぶ。そして、研究仮説から予測される結果が研究データでも観察されるか（仮説が正しいか）検証を行う、この際に用いる考え方を演繹法と呼ぶ。つまり、臨床研究は帰納法を用いて研究仮説を立て、演繹法を用いてその研究仮説の検証を行う。これがいわゆる「仮説演繹法」と呼ばれる所以である。

2. 疫学指標（epidemiologic measure）と効果指標（effect measure）

第1回でも述べたが、仮説の検証にあたり、治療や曝露の影響を「定量的」に推定することが疫学の基本的考えである。この時に必要となるものが、疫学指標（epidemiologic measure）と効果指標（effect measure）である。

疫学指標とは、治療（曝露）群やコントロール群における疾病や有害事象が存在する割合や発生頻度について示す指標であり、有病割合（prevalence proportion）、リスク（risk）、率（rate）、オッズ（odds）が基本である。以下にそれぞれの注意点を示す。

①有病割合（prevalence proportion）

有病割合（prevalence proportion）とは、「ある一時点」や「ある一定期間」における疾病や有害事象の「存在する割合」であり、「 $\text{疾病数} \div \text{全体数}$ 」で計算される。通常、英語ではprevalenceという表記で用いられる。具体例の記述としては、『2009年4月における医療従事者の調査においてうつ病の有病割合は20%であった』、というように用いられる。

②リスク（risk）

リスク（risk）とは、有病割合と混同する方もいると思うが、異なる点は「ある一時点」や「ある一定

期間」に「存在する割合」ではなく、「新たに発生する割合」を示すものである。具体例の記述としては、『くも膜下出血患者において発症後2週間以内の再出血リスクは20%である』、というように用いられる。また、細かいことであるが、日本語訳で時にみられる「発生率」という言葉は「率」と「割合」を混同した不適切な表現であるため用いない方が良い。

③率（rate）

率（rate）とは、「単位時間あたり」における疾病や有害事象の「発生頻度」を示すものであり、疾病や有害事象のもつスピードとして解釈される。罹患率（incidence rate）の場合は、人年法（人月法）を用いて「 $\text{発生疾病数} \div \text{総人年（人月）}$ 」で計算される。具体例の記述としては、『日本人男性における心筋梗塞の罹患率は30症例/10万人・年である』、というように用いられる。

④オッズ（odds）

オッズ（odds）とは、競馬のオッズと同様の概念であり、ある疾病が発生（存在）する割合ないしは確率（P）を発生（存在）しない割合ないしは確率（1-P）で割ったものである。上記の有病割合・リスクともにオッズに変換することができる。具体例としてオッズのみを表記することはほとんどなく、オッズ比として用いられる。

一方、効果指標とは、治療（曝露）群とコントロール群における疫学指標の「差」や「比」を用いて比較を行い、治療や曝露の影響を定量的に示す指標である。「差」を用いた指標の例としては、リスク差（risk difference）、率差（rate difference）が挙げられる。「比」を用いた指標の例としては、オッズ比（odds ratio）、リスク比（risk ratio）、率比（rate ratio）、ハザード比（hazard ratio）、標準化死亡比（SMR: standardized mortality ratio）などがある。相対危険度（relative risk）という用語は、リスク比、率比、ハザード比を含んだ概念として用いられることもあり注意が必要である。また、効果指標として様々なものが

存在するが、①ある条件下ではオッズ比＝率比，リスク比＝率比となること，②一般的にハザード比＝率比となることから，どれも率比の推定になっているということが疫学における理論的背景である^{1,2)}。

3. 研究デザイン

研究デザインについて基本となるのは，ランダム化比較試験 (RCT: randomized controlled trial)，コホート研究 (cohort study)，症例対照研究 (case-control study)，横断研究 (cross-sectional study) である。とくに研究仮説によって適切なデザインを選択することが重要である。

①ランダム化比較試験 (RCT: randomized controlled trial)

ランダム化比較試験 (RCT) は，曝露 (治療) をランダムに割りつけることにより，交絡要因を研究デザインで調整し，曝露 (治療) のアウトカムに対する影響を推定しようとする研究デザインである。因果関係の立証には強力であるが問題点となるのは，研究対象となった患者群の属性や追跡期間中の脱落が一般臨床の患者群と違いがないかなど，研究結果の「一般化」に関しての点である。また，研究仮説によっては倫理面や予算面から実施が難しいことが多いのも問題点である。

②コホート研究 (cohort study)

コホート研究は，時間軸において2点以上で測定を行い，曝露 (治療) のアウトカムへの影響を推定しようとする研究デザインである。治療を伴わない観察研究のみをコホート研究とみなす臨床医も多いが，それは間違いである。CPA症例において，バイスタンダーCPRが行われた群と行われなかった群の30日後の死亡リスクを検討するような研究は，まさにCPA症例を集めたコホート研究である。RCTの議論に加えて問題点としては，交絡要因の調整が中心となる。

③症例対照研究 (case-control study)

症例対照研究は，症例と対照 (コントロール群) を別々にサンプリングしてくる研究デザインであり，

コントロール群のサンプリングの仕方によって，nested case-control研究やcase-cohort研究，case-cross-over研究，前向きcase-control研究などさまざまな種類が存在する。

④横断研究 (cross-sectional study)

横断研究は，時間軸において1点のみで測定を行い，曝露 (治療) のアウトカムへの影響を推定しようとする研究デザインである。RCTやコホート研究の問題点に加えて，時間軸において1点のみでの測定になるため，因果関係の立証には弱いデザインであること，研究仮説によっては「逆の因果 (reverse causation)」と呼ばれる問題が常に残る。例えば，所得とうつ病の関係を横断調査で行った場合，所得が低いのでうつ病になりやすいのか，うつ病のために所得が低いのか，結論づけることは難しい。

+ 論文の読み方 (批判的吟味 critical appraisal) +

筆者らの経験からは，論文の書き方を知ることが論文の読み方を理解する近道である。とくにそれぞれのセクションの各段落には，どのような内容が書かれているかある程度決まっているため，その構造を理解すれば論文が非常に読みやすくなる。論文を効率的に読むには論文の段落構造を理解し，論文を段落ごとに読んでいく「パラグラフリーディング」を行えるようになることが目標である。

現在では，それぞれの研究デザインによって論文の書き方のガイドラインが出ており，非常に有用である。RCTではCONSORT⁴⁾，観察研究ではSTROBE⁵⁾，診断の精度に関する研究ではSTARD⁶⁾というそれぞれのガイドラインが出ている。以下では観察研究のガイドラインであるSTROBEを中心に観察研究の論文の読み方について紹介する。また，批判的吟味を行う上での参考書も1つ挙げておく⁷⁾。

1. Title & Abstract

批判的吟味を行うにあたり、まず行うことは論文 Title と Abstract をしっかりと読み、研究仮説・研究デザイン・概要をつかむことである。STROBEでは、論文 Title は研究デザインが一目で分かるように記述すること、Abstract は読者が研究仮説・研究デザイン・概要を理解し、読むに値するものか判断できるような記述を推奨している。また、論文 Title と Abstract は、雑誌の editor が reviewer に回すかどうかの判断に大きく影響を与える論文の最も重要なセクションであり、論文を投稿する研究者としても記述には非常に神経を使う部分である。つまり、Title と Abstract を読んで研究概要が理解できないような論文は、論文全体の質が低い場合がほとんどであり、忙しい臨床医が批判的吟味を行うに値しない論文と考えてよい。

①研究仮説の理解

研究仮説の基本はEBM関連の教科書でPE(I)COと記述されているように、

- ・ Patient: どのような対象者において
- ・ Exposure (Intervention): どのような曝露（介入）があると
- ・ Comparison (Control): どのような対照群と比べて
- ・ Outcome: アウトカム指標がどのように変わるかということが容易に理解できるものに落とし込まれていなければならない。多くの日本語論文においては、この研究仮説すらない場合や、研究仮説があったとしても十分に絞り込まれていない場合がほとんどであり、論文を読むのに非常に苦勞する。

②研究デザイン・概要の理解

そして、研究仮説の基本構造を意識しながら Title と Abstract をみてみると、最近では Title の最後に研究デザインが示されているものや、研究仮説をそのまま疑問文にしたような Title も見受けられる。ほとんどの Abstract では Background, Methods, Results, Conclusion など各セクションに構造化された形になっているため、研究仮説・デザイン・概要を意識しながら読んでいくと効率的である。

2. Introduction

STROBEでも推奨されているように、Introduction セクションは通常2-3段落構成になっており、①研究の背景、②研究の必要性・目的と研究仮説、この2つが明確にそれぞれの段落に記述されている。なお現在では、Introductionにこれまでの先行研究の膨大なレビューを記述することは避けるように推奨されている。

①研究の背景

最初の段落にはこれまでの研究概要と問題点が簡潔に論理的にまとめられており、基本的にはその後に提示する研究仮説に焦点をあてるような論理の展開になっている。

②研究の必要性・目的と研究仮説

この段落では検証する研究仮説がその必要性・目的とともに明確に提示され、これまでの研究との相違点・優位性も簡潔に書かれている。

3. Methods

STROBE では Methods セクションにおいて、読者が以下の点を理解できるように研究計画と実施状況について十分な記述を求めている。

- ・ 研究の重要点について理解できる
- ・ 妥当で信頼できる答えを出せる適切な研究方法論か判断できる
- ・ 研究計画と実施のずれが問題になる程度かどうか判断できる。

また、論文で記述されていないバイアスの検討や、研究結果を臨床現場に適応できるか判断するには、この Methods セクションを読み込まなければならない。Methods セクションは通常少なくとも、①研究デザインと研究が行われたセッティング、②研究の対象者、③測定した変数と変数の測定方法、④変数の取り扱い方やカテゴリー化、⑤統計手法、これら5項目についての記述を中心に構成されている。

①研究デザインと研究が行われたセッティング

研究結果の一般化可能性を考える上で、研究が行

われたセッティングの情報は必要である。研究の母集団についての基本情報や曝露（治療）情報についての記述や、研究を行った地域の情報、リクルート期間、フォローアップ期間、データ取得の方法の記述、既存のデータベースを利用した研究ではデータベースの特性についての記述である。

②研究の対象者

こちらも研究の一般化可能性や選択バイアスを考える上で必要な情報であり、研究仮説に合った適切な対象者の選択ができていないかに注意が必要である。対象者の組み入れや除外の基準としての選別基準（eligibility criteria）、追跡期間中の脱落例に関する記述は非常に重要である。また、研究デザイン上で効率を高めるために行うマッチング、対象者の限定についてもこの段落に記述されている。

③測定した変数と変数の測定方法

研究仮説で確認した曝露（治療）とアウトカム、交絡要因について確認を行い、それぞれがどのような方法・精度で測定されたのか確認を行う。とくにアウトカム情報についてはどのような診断基準等で測定したかを読み、信頼に足るものかを確認する。

④変数の取り扱い方やカテゴリー化

変数をカテゴリー化して結果を表示することが多いが、そのカテゴリー化が合理的なものかを判断する。とくに、効果指標を推定する際に基準となるカテゴリーが適切かどうか、無理に有意差を出すためのカテゴリー化になっていないかに注意する。また、情報バイアスを検討する際に問題となる誤分類が起きる可能性がないかにも注意する（誤分類に関しては次回以降で紹介する）。

⑤統計手法

交絡要因の調整を行うために、どのように統計手法を使ったのか、多変量解析モデルをどのように構築したのかについて理解することが必要である。種々の統計手法が存在するが、理論的にも納得できる交絡要因の選択・調整が行われているかが最も重要である。また、欠損値の扱い方や、コホート研究

における追跡不明例の扱い方も、選択バイアスや研究結果の一般化可能性を考える上で重要である。

4. Results

Resultsセクションは、①研究対象者、②記述データ、③アウトカムデータと主要結果、④その他の解析、これら4項目についての記述を中心に構成されている。基本的には客観的な記述を行い、著者の解釈や意見などは入らない。

①研究対象者

研究結果の一般化可能性と選択バイアスを検討するためにこの段落を読む。STROBEで推奨しているような、研究対象者のリクルーティングから統計解析の対象になるまでのフローチャートが掲載されていればそれを確認し、掲載されていない場合はフローチャートを頭にイメージしながら論文を読んでいく。

②記述データ

通常は論文中のTable 1として基本属性と交絡要因の分布が示されており、これを基に研究結果の一般化可能性を検討する。また、バイアスを検討するにあたり、欠損値の割合にも注意する。欠損値を示さず、研究対象者として示されている人数と基本属性の各項目の人数が異なる場合もあるので注意しなければならない。

③アウトカムデータと主要結果

研究の主要結果を示すテーブルは、批判的吟味にあたり十分に読み込まなければならない。通常はオッズ比などの効果指標が95%信頼区間とともに、粗解析（調整前）結果と多変量解析等による交絡要因の調整後の結果が合わせて提示されている。リスク比だけでなくリスク差にも変換して解釈してみることや、調整の前後の効果指標を比較して「どの程度」「どの方向に」変化したかに注目することが大切である。また、調整に用いた交絡要因の選択が合理的であるか、曝露変数のカテゴリー化は適切かなど、Methodsセクションの該当箇所を読みなおすことも

重要である。

④その他の解析

主解析とは別の解析として、性別や年齢カテゴリー、疾病の有無によりグループを分け、それぞれのグループにおいて効果指標を推定する。これをサブグループ解析といい、男女で曝露（治療）の影響が異なるのか、年齢カテゴリーで影響が異なるのかをみている。疫学用語では効果修飾（effect modification）が存在するかどうかを確認しているという。また、曝露カテゴリーを変化させて効果指標に変化がないか確認する、感度分析（sensitivity analysis）を用いて、推定された効果指標の安定性を確認することもある。

5. Discussion

DiscussionセクションはAnnals of Internal Medicine⁸⁾も推奨しているように、①主要結果のサマリー、②考えられるメカニズムと説明、③先行研究や関連する研究との比較、④研究の限界点、⑤臨床・研究への示唆、これら5項目を中心に構成されている。DiscussionセクションにIntroductionでも書かれているような先行研究の内容を繰り返すことは（日本語論文で散見される形式であるが）推奨されない。批判的吟味の際に注意する点としては、研究の妥当性や研究のもつ意味について十分に検討した記述がされているか、研究結果が過大解釈されていたり、ゆがめられた解釈になっていないかに注意しなければならない。つまり、Resultsセクションで記述された客観的データをもとにして、適切に研究結果の解釈がされているかどうか大切である。

①主要結果のサマリー

主要結果のサマリーは、基本的にはDiscussionセクションの第1段落目に書かれている。研究仮説をTitleやAbstractで把握し、主要結果のテーブルをみて、Discussionの第1段落を確認すれば、論文の主張したいことはだいたい理解できる。また、第2段落や第3段落では主要結果から導かれる解釈を記述する段落になるが、批判的吟味にあたっては、Resultsに

記述された客観的結果をもとに適切に論理が展開されているか十分な検討が必要である。

②考えられるメカニズムと説明

研究結果の解釈において、曝露とアウトカムをつなぐメカニズムについて議論することが多い。生物学的メカニズムをミクロレベルや遺伝子レベルでの知見を引用しながら検討する段落である。ただし、メカニズムの議論自体は批判的吟味に影響を与えない。

③先行研究や関連する研究との比較

研究結果の解釈にあたって、先行研究と比較しながら議論を進めていく。批判的吟味においては、著者に都合が良いように先行研究を解釈していないか、先行研究の問題点を踏まえて結果の解釈が行えているか注意が必要である。

④研究の限界点、⑤臨床・研究への示唆

研究の限界点や臨床・研究への示唆については、妥当性、一般化可能性、精度を検討することが基本である。妥当性に関しては、バイアスの方向性と大きさについての考察が必要であり、とくにバイアスの方向性については常に検討する癖をつけていただきたい。一般化可能性に関しては、研究対象者の選定と基本属性を考慮しながら判断をしなければならない。精度に関しては、95%信頼区間の幅をみるだけでなく、各変数の測定精度についても注意を払わなければならない。

今後の予定

本稿では疫学の基礎的概念について整理を行い、パラグラフリーディングによる論文の読み方について解説を行った。論文の読み方については、各種ガイドラインを参考に是非英語文献を利用して各施設での抄読会などで利用していただきたい。また、学会での発表、日本語雑誌への投稿にあたっては、本稿を参考にいただければ幸いである。次回は、研究を行う上でかなりの労力を注ぐ研究仮説の立て方と、近年の疫学分野において注目されてきている

Directed Acyclic Graph について紹介する。

第3回：研究仮説の立て方と Directed Acyclic Graph
(疫学各論2)

第4回：バイアスの考え方，結果の解釈の仕方（疫学各論3）

第5回：臨床研究における統計学の役割（疫学各論4）

参考文献

- 1) Kenneth J. Rothman: Epidemiology: An Introduction. Oxford University Press, USA, 2002.
- 2) Kenneth J. Rothman, Sander Greenland, Timothy L. Lash: Modern Epidemiology 3rd ed. Lippincott Williams & Wilkins, USA, 2007.
- 3) Miguel Porta: A Dictionary of Epidemiology 5th ed. Oxford University Press, USA, 2008.
- 4) The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration (Ann Intern Med 2001; 134: 663-94)
<http://www.consort-statement.org/index.aspx?o=1011#E&E>
- 5) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration (Ann Intern Med 2007; 147 (8): W163-94)
<http://www.annals.org/cgi/content/full/147/8/W-163>
- 6) The STARD Statement for Reporting Studies of Diagnostic Accuracy: Explanation and Elaboration (Ann Intern Med 2003; 138: W1-12)
<http://www.stard-statement.org/>
- 7) J. Mark Elwood: Critical Appraisal of Epidemiological Studies & Clinical Trials. Oxford University Press, 2007.
- 8) Annals of Internal Medicine: Information for Authors.
http://www.annals.org/shared/author_info.html#overview

ABSTRACT

Short Seminars on Epidemiology for Clinician

Clinical Research Based on Community Hospitals

Lesson 2: Basic Knowledge of Epidemiology and How to Read Paper

Hirokazu Komatsu ^{a,b}, Etsuji Suzuki ^b, Hiroyuki Doi ^b

^a Division of Community Care, Saku Central Hospital,
Nagano Prefectural Federation of Agricultural Cooperatives for Health and Welfare

^b Department of Epidemiology, Okayama University Graduate School of Medicine,
Dentistry and Pharmaceutical Sciences

Comprehending academic theses (i.e., critical appraisal) is a fundamental technique that is necessary when conducting research: in recent days with the practice of EBM, it has become a prerequisite needed even by clinicians. A great many clinicians find understanding academic theses almost painful. However, the authors believe that acquiring a basic knowledge of epidemiology and studying the methods of writing academic theses are the best approach if one wishes to master a technique that will lead to an understanding of academic theses. For the fundamental knowledge of epidemiology, understanding epidemiological indices and outcome measures, which are necessary for quantitative estimation of the effects of exposure (treatment) on the outcome, as well as the research design that are a prerequisite in considering biases are extremely important. For an interpretation of academic theses, on the other hand, one's efficacy to read and understand academic theses improves dramatically if he understands the paragraph structure and learns to do "paragraph reading" (reading each paragraph consecutively). The best training is to read these academic theses while referring to various guidelines for writing papers.

(JJAAM 2009; 20: 338-44)

Key words: methods to read and understand academic theses, epidemiological indices, outcome measures, research design, paragraph reading