

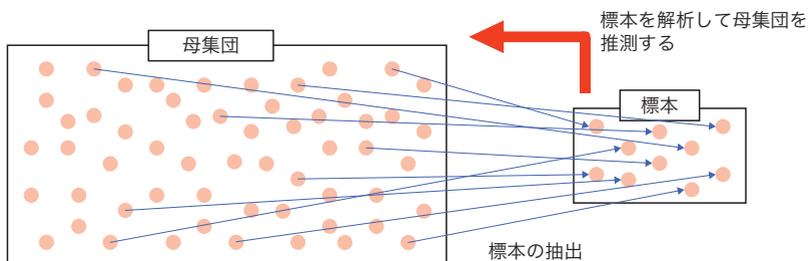
## 1 なぜ統計解析が必要なのか？

人間は自分自身の経験に基づいて、感覚的にものごとを判断しがちである。例えばある疾患に対する標準治療薬の有効率が50%であったとする。そこに新薬が登場し、ある医師がその新薬を5人の患者に使ったところ、4人が有効と判定されたとしたら、多くの医師はこれまでの標準治療薬よりも新薬のほうが有効性が高そうだと感じることだろう。しかし、たまたま有効性が出やすい5人に治療が行われたにすぎないかもしれない。同じ疾患を有する患者であったとしても、疾患の細かな分類や進行度、患者の年齢、性別、臓器の状態などによって有効率は左右される。さらに背景の条件が全く同じであったとしても有効率にばらつきは生じる。治療に対して思い入れが強ければ強いほど、治療結果に大きく一喜一憂し、客観的な評価が困難となる。印象に残る結果は感覚的な判断を偏らせてしまう。

**統計解析の目的**は、前提としてこのような様々なばらつきが存在する状況の中で、限られた**標本 (sample)** から**母集団 (population)** を**推測**し、より一般的な結論を導き出そうとすることである。**母集団**の定義は状況によって異なるが、例えばある疾患に対する新薬の有効性を評価するのであれば、その疾患を有する全ての患者が母集団となる。統計解析をしていると、目の前にあるデータだけを対象としているような錯覚にとらわれることがあるが、実際に行っていることは、その標本を用いて本当の**母集団の全体像を推定**しようとしているのである（選挙の出口調査による全体の投票数や議席数の予測をイメージすればよい）。



**統計解析の目的は、母集団から抽出した標本（サンプル）を用いて解析することによって、母集団を推測することである**



## 2 変数の種類とその要約

### 1 変数の種類

統計解析で扱う主な変数は**連続変数 (continuous variable)**、**順序変数 (ordinal variable)**、**名義変数 (nominal variable)** の3つに分けられる。連続変数は身長、体重など、数値で表される**定量的なデータ**を意味する。順序変数、名義変数はいずれも**質的なデータ (カテゴリー変数、categorical variable)**であるが、順序変数は尿蛋白の(-)、(±)、(+)、(2+)、(3+) や、腫瘍の進行度のステージ I、II、III、IVのように順序づけられたものである。一方、名義変数は、性別の男性、女性や、ABO血液型のA、B、O、AB型のように順序の関係がない(男性、女性、あるいは有効、無効のように二値だけを持つ場合は**二値変数**あるいは**二区分変数 (binary variable)**とも呼ばれる)。

特殊な変数として医学統計ではしばしば**生存期間**の解析が行われる。正確にいうと、必ずしも生存期間だけを対象とする解析ではなく、ある時点からあるできごと(イベント)が発生するまでの期間(**time-to-event variable**)の解析であり、死亡がイベントとして定義された場合に生存期間の解析が行われることになる。この解析方法の特徴は、ある時期まで生存していた(あるいはイベントが発生していなかった)ことは知られているが、その後の情報が得られないような場合に**観察打ち切り (censor)**として解析に含めることができる点である。例えば、ある疾患に対して特定の治療を行った後の生存期間を解析する場合に、最終観察時点で生存中の患者の真の生存期間は不明であるが、その時点で打ち切りとして扱うことによって解析に含めることができる。この解析においては、イベントが1回しか発生しないものであることと、打ち切りとなる理由が解析対象のイベントの発生とは無関係であることが必要である。例えば悪性腫瘍に対する化学療法後の生存期間の解析において、打ち切りとなった理由が他院への転院のような場合は、病状が増悪して死期の近づいた患者がしばしばホスピスに転院するという背景が解析上の**偏り (バイアス)**を生じてしまう危険性がある。

### 2 変数の要約、信頼区間

各変数を**要約**して記述する方法はそれぞれの解析のところで詳しく述べるが、まずは全体像を眺めることが重要である。名義変数なら**頻度分布**を、連続変数であれば**散布図**、**ヒストグラム**、**箱ひげ図**などを描いてみる。生存期間を表すためには**Kaplan-Meier 曲線**が用いられる。各変数を端的に記述するには、そ

影響を考慮して  $P < 0.10 \sim 0.20$  のように広めに選択する) を全てモデルに含めるという方法である。これらを全て含んだモデルを最終モデルとすることもありますが、変数を限定する場合には、機械的に  $P \geq 0.05$  の変数の中から  $P$  値が最も高い変数を1つずつ除外して再計算することを繰り返し、最終的に有意な ( $P < 0.05$  の) 変数だけを残す方法 (**backward stepwise deletion**) や、モデルの適合度のよさの指標 ( $F$  検定、尤度比検定、AIC、BIC など) をみながら変数を増減することによって、最も当てはまりがよいモデルを決定する方法などが行われる。ただし、ある特定の変数に関心があるような場合は、その変数は常にモデルに残す必要がある。理論的に、あるいは過去の研究結果などによって明らかに重要である因子も同様に除外しないようにする。

## 8 EBM と臨床研究

1990年代から **EBM (evidence-based medicine)** という用語がもてはやされるようになった。EBMは**臨床疫学 (clinical epidemiology)** を臨床現場の診療により近づけたものであり、個々の患者の治療を決定する際に、その時点での最良の**エビデンス** (臨床研究の実証結果) を解釈したうえで、患者の状態や価値観をふまえて慎重に適用することである。EBMの実践はステップ1「問題点を定式化する」、ステップ2「その質問に答えるためのエビデンスを収集する」、ステップ3「得られたエビデンスの信頼性を批判的に吟味する」、ステップ4「得られた情報が目の前の患者にあてはめられるかどうかを検討す



### クリニカルクエスションの構成要素

P	Patients	対象とする患者
I (E)	Intervention (Exposure)	介入や曝露 (治療、検査など)
C	Control	比較する対照群
O	Outcome	結果、転帰 (治療効果、生存率など)



### EBMの5つのステップ

ステップ1	患者の問題の定式化
ステップ2	問題についてのエビデンスの収集
ステップ3	エビデンスの批判的吟味
ステップ4	エビデンスの患者への適応
ステップ5	1~4のステップの評価

る」、ステップ5「自分が行った1から4のステップを自己評価する」に沿って行われる。

ステップ1では臨床現場で直面する疑問を明確な言葉（**クリニカルクエスチョン**）に置き換える作業であり、「どのような患者に、どのような治療を行ったら、対照群と比較して、どのような結果の違いが生じるか」という文章で表す。構成要素のそれぞれの頭文字をとって **PICO** あるいは **PECO** と呼ばれる（表）。一般に観察研究では何らかの要因への曝露（**exposure**）の影響を評価するので **PECO** となり、介入研究では何らかの介入（**intervention**）の影響を評価するので **PICO** となる。

ステップ2では **PICO** で得られたキーワードを利用して文献検索を行う。網羅的な検索を行うにはインターネットで **PubMed** などの**一次資料**（主として専門誌に掲載された学術論文）を利用することができるが、数多くの文献が抽出された場合にその中から重要なものを見極めていくにはある程度の知識が必要である。そこで、不慣れな分野で検索を行う場合は、まずはエビデンスを集積した**二次資料**から読み始めることを推奨する。例えば専門雑誌の**総説**（**review article**）や **UpToDate**、**Clinical Evidence**、**Cochrane library** なども有用である。

ステップ3では得られたエビデンスの信頼性、妥当性について客観的な評価を行う。このステップでは基本的に各研究の中で行われた解析、結論が**バイアス**や偶然によって歪められていないかどうかを科学的に検証する（すなわち**内的妥当性**の評価である）。また、結果の臨床的な重要性についても検討しなければならない。有意差が得られたということと臨床的な重要性（効果の大きさ、**エフェクトサイズ**）とは別問題である。有用な二次資料はこのステップ2、3を集約してくれる場合もある。

そして、ステップ1~3で得られたエビデンスが目の前の患者に適用できるかどうか（**外的妥当性**の評価）を検討するのがステップ4である。臨床試験の多くは年齢、疾患の進行度、合併症など、様々な要素で適格条件を絞りこんでいるため、その試験で得られた結果が一般的な患者に応用できるとは限らないのである。患者や家族の人生観や社会環境なども治療決断に組み込まなければならない。

最後にステップ1~4で行ったことをステップ5で評価し、次の機会に役立つ。

これらのステップ（特にステップ4）をみればわかるとおり、EBMとは画一的な料理帳のような治療を強要するものではない（David L Sackett は

Evidence-based medicine is not “cook-book” medicine. と記述している)。数多くの診療ガイドラインが出版されているが、これらについてもガイドラインに従っていればよい、あるいは従わなければならないというような金科玉条ではない。ガイドラインはあくまで「一般的な」診療を示すものにすぎず、それが該当する患者はおそらく60~80%程度であろう。また、決してメタアナリシスや大規模無作為割付比較試験によって生み出された、いわゆる Level I エビデンスだけが求められるわけでもない。現時点で存在するエビデンスをいかにして個々の患者の診療に役立てるかに医師の力量が問われる。

一方、日常診療においてエビデンスを検索すると、十分なエビデンスが見つからないことが多い。そのため、よりよいエビデンスを生み出すための臨床研究が必要になる。すなわち EBM も臨床研究も、いずれも日常診療において生じる疑問から始まる表裏一体の関係にある。



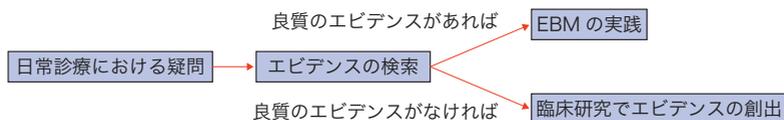
### エビデンスレベルの分類の例

(厚生労働科学研究費のガイドラインの作成の手順より)

- |      |                            |
|------|----------------------------|
| I.   | システマティックレビュー/メタアナリシス       |
| II.  | 1つ以上のランダム化比較試験による          |
| III. | 非ランダム化比較試験による              |
| IV.  | 分析疫学的研究(コホート研究や症例対照研究による)  |
| V.   | 記述研究(症例報告やケース・シリーズ)による     |
| VI.  | 患者データに基づかない、専門委員会や専門家個人の意見 |



### EBM と臨床研究の関係



## 9 観察研究と介入研究

臨床研究は**観察研究**と**介入研究**に大きく分類される。介入研究は研究者が被験者に対して何らかの**介入**(治療、検査、割付など)を行う研究であり、必然的に**前向き(前方視的、prospective)**なデザインとなる。一方、観察研究は研究者が被験者に対して介入を行わずに、ある因子への曝露と疾患の発症の関

## 16 臨床研究の倫理

第二次大戦中にナチスが強制収容所のユダヤ人に対して人体実験を行ったことはよく知られているが、大戦以降にも**非倫理的な臨床試験**は数多く行われた。米国では1945年から1965年の間にMassachusetts General Hospital、National Institutes of Healthの研究費はそれぞれ17倍、624倍に増加し、それにともない倫理的問題は数的にも質的にも増加した。Beecherは1964年に有名誌に掲載された100件の臨床試験を調査し、12件は非倫理的であると考えた。22の実例が記されているが、最も印象的な一例を示す。「悪性黒色腫の患者から、腫瘍細胞がその母親に移植された。その目的は腫瘍免疫の解明と、腫瘍に対する抗体が産生されれば患者の治療に役立つかもしれないという望みである。しかし既に終末期であった患者（娘）は移植の翌日に死亡したため、後者の目的は現実的ではなかったであろう。母親に移植した腫瘍細胞は移植24日目に広く切除されたが、母親は移植後451日目に悪性黒色腫の全身転移で死亡した。」研究者の真の意図は不明であるが、研究者がいかに患者の治療を考えていたとしても、そして母親から**インフォームド・コンセント**を得ていたとしても、許されることのない臨床試験であることは明らかである。そして、本例のように、親族が治療の対象となっている場合、家族から適切なインフォームド・コンセントを得ることが困難であることも容易に想像される。

前述のナチスの人体実験はニュールンベルクの国際軍事裁判所で裁かれた。その判決文の中で臨床試験の基本原則として10項目からなる基本理念が制定された。これが**ニュールンベルク綱領**（1947年）であり、医学的な実験には「被験者の自発的な承諾が前提であり、決定を下すことができる知識が与えられ、十分な理解を得た上で、被験者本人の同意が必要である」というインフォームド・コンセントの概念が記されている。**世界医師会**はこれに基づき、1948年に**ジュネーブ宣言**、1949年に**医の倫理の国際綱領**、1954年に**研究・実験の原則**を発表し、さらに、1964年に人を対象とする医学研究の倫理的原則として**ヘルシンキ宣言**を発表した。その後、1975年、1983年、1989年、1996年、2000年、2008年、2013年に一部修正されているが、現在でも臨床試験の倫理的原則として最も重要なものであり、臨床試験に携わる全ての者が一読しておく必要がある。2013年版の全文の和訳も様々なホームページに掲載されている。国内の臨床研究の倫理指針としては、「疫学研究に関する倫理指針」と「臨床研究に関する倫理指針」が統合され、新たに「**人を対象とする医学系研究に関する倫理指針**」が制定された。

倫理的原則に従って作成したプロトコールは、研究開始前に**倫理委員会**の審査を受けなければならない。後方視的研究も倫理委員会での審査が必要である。また、有意な結果が得られなかった臨床試験が公表されないという**出版バイアス**の増加を避けるため、臨床試験は実施前に登録することが推奨されており、既に国際的な主要な医学誌はあらかじめ登録されている臨床試験以外は掲載を認めていない。国内では**UMIN-CTR** (<http://www.umin.ac.jp/ctr/index-j.htm>) に登録することができる。

また、全ての臨床試験は被験者から文書で同意を得る必要がある。後方視的な観察研究については必ずしも個々の被験者からの同意を得なくてもよいとされているが、ホームページなどで研究の情報については公開する必要がある。

## 17 臨床試験の報告に必要な要素

臨床試験の報告方法について、いくつかのガイドラインが提唱されている。特に**無作為割付比較試験**は**CONSORT (Consolidated Standards of Reporting Trials)**に従って報告することが求められる。当然ながら研究計画の段階で CONSORT に含まれる要素は入念に検討されていなくてはならない。CONSORT 声明の原文は <http://www.consort-statement.org/> で、日本語版のチェックリストは [http://www.lifescience.co.jp/yk/jpt\\_online/index\\_jpt.html](http://www.lifescience.co.jp/yk/jpt_online/index_jpt.html) に公開されている。