

【臨床研究のストラテジー】

## 臨床研究のストラテジー

### 第1回：臨床研究のエレメント

東京慈恵会医科大学臨床研究開発室

浦島 充佳      松島 雅人

#### まえがき

臨床研究とは、より効果的に診断・治療する、あるいは、病気を予防するために、臨床の場から疫学的手法を用いてエビデンスを構築する学問である。そして得られたエビデンスは医療や予防医学を介して社会に還元されて意味をもつ。すなわち、臨床研究はあくまで‘病人を治す’および‘疾患発生を予防する’ための学問であって実験研究のように‘病気を理解する’学問とは一線を画する。東京慈恵会医科大学学祖の高木兼寛は、ビタミンB<sub>1</sub>欠乏という脚気の原因が判明する前に脚気を撲滅してしまった。これこそが臨床研究の真髄である。「喫煙が肺癌を増やす」事実は病院内でのインタビューからはじまった。いまだに喫煙が肺癌を発症させるメカニズムの詳細は明らかにされていないが、喫煙率の低下により米国では肺癌死亡が減少に転じた。レント博士がサリドマイドの催奇形性に気付いたきっかけは犠牲者の母親の言葉であった。サリドマイド催奇形のメカニズムは解き明かされないままであるが、サリドマイド児の出生は無くなった。サリドマイドが世界各地で使われはじめたにもかかわらず、アメリカFDAがサリドマイドの臨床使用を認めなかったのは「臨床使用に耐えるエビデンスが無い」という毅然としたものであった。「牛痘に罹患した子供は生涯天然痘に罹患することはない」というエビデンスを知っていた医師ジェンナーが天然痘ウイルスの存在を知らずとも天然痘ワクチンの発想に至ったのは極自然の成り行きであるとも考えられる。このように、臨床研究は、人類が時代時代に直面した大きな問題をなんとか回避し、多くの患者を救うために多大な貢献をしてきたと言える。

2001年9月より、総合医科学研究センター内に臨床研究開発室が発足した。その使命は、慈恵大学内の臨床研究を推進支援することにある。その方法には、臨床医の仮説に対してこれを証明するための臨床研究のデザイン構築を手伝うとともに、臨床疫学の一般的方法論をセミナーや紙面を通じて広めるなどが考えられる。そこで、臨床研究のストラテジーを6回に分けて慈恵医大誌に掲載させていただくことになった。

第1回：臨床研究のエレメント

第2回：ケース・コントロール研究

第3回：コホート研究

第4回：臨床試験

第5回：診断に関する研究

第6回：遺伝疫学，感染症疫学

医学教育で有名なウィリアム・オスラー先生は、医学を「知識の伝授」とはせず、「問題を深く掘り下げて考えることのできるまで、方法論を自分のものとする」と述べている。プラトンは「教育とは、一生にわたる過程であり、学生は大学時代その第一歩を歩みだすに過ぎない」と前提した上で、「教育とは諸原理を1つ1つ教え込み、学生を正しい道に導き、方法を授け、勉強のやり方を教え、本質的なものとそうでないものを早くから識別しうる方法を教えること」と定義している。これから解説する臨床研究とはあくまで方法論（ストラテジー）であり、それに何をあてはめるかは読者次第である。

一方、臨床研究の方法論に関して中途半端な理解はマイナス面さえ有する。作家マーク・トウェインは、「普通の嘘」「黙っている嘘」「統計学」をもって「嘘には三種類ある」と指摘している。つまり、臨床研究を誤って実践すると、もっともらしく嘘

をついて、かえって患者を危険に曝す結果となってしまうかもしれない。よって、臨床研究の方法論を正しく理解することが、その実践の第一歩となる。

臨床に携わる人が、臨床研究の基本原則を理解し、本連載を通じて複雑な臨床から病める人々を救う新しい真実を汲み上げる一助となってくれることを願う次第である。

### 臨床研究のエレメント

臨床研究の方法論を理解する上で、まず知っておいていただきたい基本用語（エレメント）を解説する。実際のところ、この部分だけでも1冊の本になるような内容であるが、要点だけに絞ってまとめてみた。

#### 1. *Exposure vs. Outcome* (暴露と結果)

人は何らかの因子に暴露 (exposure) されることによって疾病などの結果 (outcome) を発生する。臨床研究の最終的な目的は、この2つの間に真の因果関係を証明することである。Exposure というと放射線被曝のようなものを思い浮かべるが、治療であっても、運動であっても、性別であってもかまわない。一方、結果というと白血病などの病気発生を思い浮かべるが、病気の再発、薬の副作用、あるいは離婚などでも outcome にし得る。

#### 2. *Discrete vs. continuous* (変数の尺度)

Discrete とは Yes/No (or 1/0) で表現されるものである。たとえば、性別、白血病である・ないなどである。一方、continuous は連続的な数値で表されるもので、白血球数や脈拍数などである。しかし、白血球数が12,000以上の時に1とし、それ未満を0とすれば、discrete data に変換することができる。臨床研究では結果を discrete な変数として測定し評価することが多い。一方、結果が連続変数である場合には t-test などと比較する。さらに、白血病初診時の白血球数を、5万以上を1、2-5万を2、2万以下を3に分けるような場合は順序を持つカテゴリ変数と解釈できる。さらに、赤、青、黄色のような順序を持たないカテゴリもある。その場合、個々の色に対して「赤である、

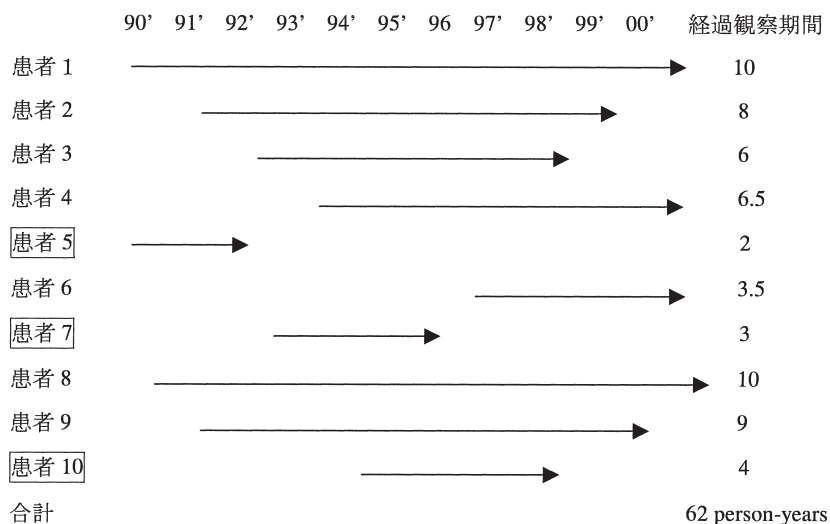
赤でない」のように分類する (dummy variable)。

#### 3. *Prevalence (罹患率) vs. Cumulative Incidence (発症率)*

Prevalence は、ある時点あるいは短い期間におけるある疾患の割合 (proportion) を指す。たとえば「日本における今週のインフルエンザ罹患率は2%であった。」のような表現がそうである。Prevalence はしばしばアンケート調査などの cross sectional study に応用される。一方、臨床研究における risk とは、ある人がある一定期間にある病気 (結果) になる確率を指し、cumulative incidence とはある集団における平均の risk を指す。例えば「3歳児が今後10年間に不慮の事故で死亡する平均のリスクは0.1%である」「ステージIの悪性リンパ腫が5年間に再発する平均のリスクは35%である。」などである。疾患発生の原因を探るためには、その疾病発生のリスクを測定することが前提となる。

Prevalence と risk/cumulative incidence を比較すると前者は point で疾患の割合を見ているのに対して、後者は一定期間にどれくらいの人が病気に罹患したかを測定している。Prevalence = Incidence × Duration としても置き換えられる。仮に感冒の prevalence が10%であり、糖尿病の prevalence が10%であった場合、前者の罹病期間 (duration) は短く、後者は長いわけで、逆に incidence は前者で高く、後者で低いことになる。つまり、prevalence には、その疾患の平均のリスクだけではなく、罹病期間も関わってくるため、病因探求には向かない (公衆衛生的には重要なファクターである)。

Cumulative incidence を考える際もう1つ重要な点は、ある集団で一定期間観察した後病気になる確率であるから、Cumulative incidence を測定する研究を開始する際、対象となる集団は病気を持っていない点を銘記しておく必要がある。逆に、病気になりえない人も除外する。早期胃癌で胃を全摘した人や、子宮筋腫で子宮を全摘した人は、胃癌や子宮癌の cumulative incidence を測定するための臨床研究対象からは外される。さらに、outcome を発生した時点で経過観察は中止される。何故ならすでに結果を発生する risk を負って



はいないからである。

#### 4. Cumulative incidence vs. Incidence rate

Cumulative incidence はある一定期間に outcome を発生する平均のリスクである。しかしながら、研究で設定した期間、すべての対象者を完全に観察し、cumulative incidence を算出することは困難である。途中から研究に加えたい場合や (variable time at entry)、途中で経過を追えなくなる場合 (loss to follow-up) など臨床研究の際しばしば経験する。このような場合、1年などの単位あたりの観察期間を設定して、下記のように person-time でとらえると問題は解決される。

3人の患児 (5, 7, 10) が白血病を再発して死亡したとすると、incidence rate =  $3/62$  person-years になる。分子が 1-100 の間に来よう分母を 10, 100, ..., 10,000 など示すのが慣例である。上の例は  $48.4/1,000$  person-years と書く方が一般的である。このような場合においてのみ rate を使用する。そして rate は person-time の単位をもつ。

#### 5. Intervention vs. Observational study

研究者側が Exposure を人為的にコントロールする研究デザインを intervention study と呼び、exposure を自然の摂理にまかせて人為的にコントロールせず観察するだけのものを observation study と呼ぶ。

intervention study の典型的なものは random-

ized controlled trial (RCT) である。しかし、同じ病院での 5 年前の成績と比較するなどの historical control の場合も intervention study に含まれる。学祖高木兼寛が同じ航路をとる軍艦で行った脚気予防食の臨床研究も intervention study の historical control study に相当する。あるいは crossover design とみなすこともできるかもしれない。また、intervention study は prospective cohort study の特殊形と考えることもできる。

#### 6. Longitudinal vs. cross sectional

Longitudinal とは exposure と outcome の間に時間差があることを示している。Exposure がきっかけで outcome が発生するとしたら、必ず exposure は outcome に先行する。よって、因果関係を捉える場合には、この時間関係を知る必要がでてくる。通常、intervention study, case-control study, cohort study (次回以降解説する) は longitudinal である。その時間は、妊娠中の薬剤がきっかけで 20 年後に腫瘍を発生する場合もあれば、目の前で車にはねられ即死する場合もあるであろう。時間差が短い方が、その因果関係をつかみやすいのは事実である。なぜなら、時間が経つほどその他、諸々の要素が加わってくるからである。一方、cross sectional study からは時間的關係をつかむことができない。たとえばアンケート調査で、子供の日頃の運動量と肥満度を調べると

する。どうやら、肥満度と運動量の間には相関関係がありそうである。しかし、運動しないから肥満になるのか、肥満だから運動しないのかは、あるいは両者の相互作用なのかは、cross sectional study からは判定できない。

### 7. Retrospective vs. prospective study

Prospective study は outcome が発生する前に観察を開始する研究で、exposure はすでにある場合もあれば、まだ exposure されていない場合もある。逆に retrospective study は outcome が発生してから研究をするため、必然的に exposure はすでに起こっている。Cohort study=prospective, case-control study=retrospective と勘違いしている人がいるが、必ずしもそうではない。

あなたは、薬剤 A の服用が潰瘍性大腸炎の発生と関連しているのではないかと仮説をたてた。たとえばあなたの病院に 12 歳の女兒が血便精査で入院し、潰瘍性大腸炎の診断がついたとする。そこで、外来カルテの中から同じ 12 歳女兒のカルテをランダムに選んでくる。そして、最近 1 年間の服薬歴を比較する。また潰瘍性大腸炎の新患があったら同じことを繰り返す。これは outcome (潰瘍性大腸炎) を軸にした研究であり、outcome が発生する前から研究をはじめるので Prospective case-control study といえる。Prospective の方が retrospective より recall bias や missing data を少なく効率的に収集できるため、retrospective case control study の欠点を補うことができる。同仮説を証明するために prospective cohort study のデザインを組むとすると、相当多人数を長期間経過観察しなくてはならず、費用もかかる。結果発生数が少ない場合、case-control study は圧倒的に有利である。

### 8. Factual outcome vs. counterfactual outcome

もしも、ある地区で放射能漏れ事故があり、地域住人が暴露されたとする。そこで、発生する白血病すべてはその放射能漏れ事故が原因なのだろうか？我々は、“その地区で同期間、放射能漏れが無かったら白血病がどれくらい発生するか？”という疑問に答える必要がある。たとえば重いもの

を持ったときより腰痛を訴えはじめた患者さんが受診し、あなたは MRI にて脊椎の disk を観察したとする。そのような目でみると disk が少し変形して脊髄の方にはみだしているようにも見える。そこであなたはこの患者さんを‘椎間板ヘルニア’と診断した。しかし、disk の変形が腰痛の原因であると判定するためにはこの患者さんの腰痛直前の MRI 所見と比較しなくてはならない。

このようにある因子の存在下で起こった事実を factual outcome, その因子が存在しなかったときに起こる事実を counterfactual outcome と呼ぶ。Exposure と outcome 間の因果関係を証明するためには正しい counterfactual outcome と比較する必要がある。たとえばある女性が子供時代側彎症に罹患して、これに対して胸部レントゲン写真撮影を頻回に受け、25 年後その女性は乳癌になったとする。さて子供時代頻回に撮影した胸部レントゲン写真が乳癌の原因なのだろうか？これを証明するためにはもう一度子供時代にかえて胸部レントゲン写真を撮らず、それ以外はまったく同じ生活をする (conditional counterfactual outcome)。そして 25 年後に乳癌にならなければそれが原因であったと断言できる。これを知るには“タイムマシン”が必要であるが、SF 小説ではないので実行不可能である。それでは、側彎症で胸部レントゲンを撮らなかった女性 1 人と比較すれば済む問題であろうか？その女性はたまたま乳癌になりやすい体質だったかもしれない。

もし多人数に対してランダムに胸部レントゲンを頻回に施行したらどうなるだろうか？ランダムにレントゲン群、非レントゲン群に分けているのでレントゲン以外の条件は集団としてほとんど一致するはずである。そして結果に何らかの違いがみられたら、それはレントゲン写真撮影 (exposure) の違いと推測できる。しかし、現実問題として、レントゲン写真を意味もなく頻回に撮影するわけにはいかない。そこで、乳癌の患者と乳癌でない人々の間で、過去に撮影した胸部レントゲン写真の撮影頻度を比較する方法をとる。逆に、ある病院で胸部レントゲン写真撮影回数によって乳癌発生頻度を比較する方法もあるであろう。しかし、将来発生する乳癌の頻度が必ずしも多くないことから、前者の方が現実的である。この counter

factorial outcome を原点にすると臨床研究のデザインを理解しやすい。

以上，臨床研究の際に重要なエレメントについて解説した。