

を引き起こした。治験受け入れ側では、大学の（少なくとも一時的な）地盤沈下が発生し、一方で治験を収益源とする開業医や Site Management Organization (SMO) の雨後の竹の子状の設立と淘汰、Contract Research Organization (CRO) の勃興、そして Clinical Research Coordinator (CRC) が試験成功のキーパーソンとして認知されるに至った。上記の EBM の考え方が普及し、日本からエビデンス発信が希薄であったことの反省とともに、医師主導臨床試験の重要性が認識され、(いまだに十分ではないものの) 厚生労働省からの臨床試験に対する研究投資の増加とともに数多くの医師主導試験実施のための組織が誕生した。

このような流れの中で臨床試験あるいは広く臨床研究に対する統計家の貢献の重要性がようやく認識され、実務家・研究者の需要が立ち上がるとともに教育システムも生まれてきた。大学では 1992 年に東京大学に初めての生物統計学の講座が誕生し、1999 年の北里大学大学院薬学研究科、2000 年の京都大学大学院医学研究科と講座開設が続いた。2002 年には東京理科大学大学院(経営工学)に社会人対象修士コースが生まれ、2004 年には久留米大学、国立保健医療科学院にも教育コースが開設された。本講演では、歴史的な展開も含め、詳細な数理というよりは、このような専門家を必要としてきた背景の記述と主要な概念、そして将来の展望について述べてみたい。

## 2. 新 GCP 以前

### 2.1 Exploratory 対 confirmatory, explanatory 対 pragmatic

臨床試験の分類機軸に exploratory (探索的) 対 confirmatory (検証的)、explanatory (説明的) 対 pragmatic (実践的) がある。これらの言葉を使えば、新 GCP 以前の日本の治験の問題は confirmatory であるべき統計解析と解釈を exploratory に行い、本来相反する explanatory と pragmatic アプローチを混同し、どちらつかずの臨床試験を行っていたことに帰着するのではなからうか？

Exploratory という言葉は、統計学全般の中では Tukey の「革命的」著書 Exploratory Data Analysis (Tukey, 1977) によっていわば提唱された。Tukey は数学的に体系化された仮説検定論 (Neyman-Pearson 理論) に代表されるこれまでの統計的データ解析を confirmatory と呼び、

- データ解析は数学的最適理論ではなく、データが語ろうとしていることをいかに抽出するかが本質
- そのためにコンピュータ利用を含む新しい手法が必要
- モデルが真実を反映しているかの診断が重要
- モデルあるいは前提と真実とが乖離していても大きな影響を受けないロバストな手法が有用

という現在の薬効評価の統計学にも通じるデータ解析の基本的哲学を提唱した。

臨床試験の分野では、confirmatory と exploratory の区別は、(臨床試験に携わる統計家にとって周知であったが) ICH-E9 統計ガイドライン (厚生労働省, 1998; ICH E9 EWG, 1999) により強く認識されるにいたった。もちろん一つの臨床試験にはいずれの側面も存在し、統計解析計画書では検証的部分と探索的部分を明確に分けることが要求されている。より後期の試験、たとえば第 III 相試験では検証的側面が主になるものの、有効なサブグループの探索 (治療効果と背景要因の交互作用の検討) など探索的解析も適正使用のために重要な情報を提供する。

同ガイドラインにより、多重性 (multiplicity)、すなわち検定の繰り返しによって第一種の過誤を上昇させる行為はきつく戒められることとなり、これは現在のわが国の審査にも厳密に適用されている。多重エンドポイント・多重時点・多重サブグループ・多群の場合の2群比較の繰り返し・複数の検定方法がこの例である。割り付け開示後の追加解析も広義にはこれに含まれる。この分類機軸に対する理解は、研究者 (スポンサー) 側・審査側にも十分浸透していると思われる。

なお、臨床試験における confirmation の統計的・数学的な基盤は、ランダム化で生成される確率空間から計算される p 値である。すべての応用分野を通じ p 値の数学的根拠が明確な場は二つしか存在しない。無作為抽出 (サンプリング) あるいはランダム化がなされている場面である。前者が実質不可能な臨床試験においては、結論の一般化可能性を当面は犠牲としても confirmation のためにランダム化を行うのである。したがって、確率空間あるいは結論をゆがめる (バイアスを生む、あるいは第一種の過誤を増大させる) 解析上の取り扱いは極力避けることになる。そのために採用されるのが Intent-To-Treat (以下略して ITT) の考え方である。

10 あまりの臨床試験に関する代表的な教科書を渉猟してみたものの、explanatory と pragmatic の区別について詳しい記載をしているものは以外に少ない。最も詳しいものが全面的に両者の違いを押し出している特長のある教科書 Schwartz, Flamant and Lellouch (1980) であり、Piantadosi (1997) が1ページほど、Pocock の著名な教科書 (原著は 1983) ではプロトコル不遵守例の解析上の取り扱いに、Redmond and Colton の辞典 (2001) では欠損データの扱いのところに登場するのみである。がんを中心とした教科書に記載が見当たらないのは、がん領域の比較試験では pragmatic アプローチが当然であることの反映かもしれない。それぞれの表現を、私見も交えてまとめてみると表 1 のようになる。もちろんこの分類は相対的・哲学的なものであり、一つの試験の性格付けをこの規準から常に明確にできるとは限らないが、

表 1 Exploratory 対 confirmatory, explanatory 対 pragmatic

Explanatory	Pragmatic	
目的は明確で広く理解されている	科学者にはよく理解されない	Flamant et al.
実験室での実験の拡張	臨床実践への指針を与える	Flamant et al.
仮説検定	治療方針選択	Flamant et al.
method-effectiveness	use-effectiveness	Piantadosi
両者は compete する		Piantadosi
analysis of compliers only	analysis by inten(tion)-to-treat	Pocock
PPS (Per-Protocol-Set)	FAS (Full-Analysis-Set)	ICH E9
欠損での処理は what if ? で	total impact を評価	Redmond&Colton
proof of concept 試験	simple and large-scale evidence	

## 2.2 これまでの反省

上記の分類機軸を用いれば新 GCP 施行前の、わが国の臨床試験の方法論上の問題点は以下のようである（詳しくは大橋, 2004）:

0. Confirmation という概念の理解が申請者側・審査側とも不十分であった
  1. 1 施設あたりの症例数が少ない
  2. 実薬を対照とした同等性試験で最終的な検証が行われる
  3. 主治医による主観的総合評価が主たる（プライマリ）エンドポイントである

0 については既に述べた。全体で有意性が証明できない場合にサブグループの結果から承認にいたった例もあったように思われるが、(2 に関連して) 有意でない場合には同等として承認されたという最大の問題がある。これが 2 の実薬比較試験の感度の問題 (assay sensitivity) も合わせ、ホパテ市場撤退後の抗痴呆薬評価において、多くの薬剤がプラセボ対照試験で失敗するという惨憺たる結果の背景となった。

稀少疾患については、1 施設の登録例数が少ないことは当然である。このような場合でもランダム化が正しく行われれば結論の妥当性は保証され、かつ施設が多いことから試験は pragmatic で一般化可能性も高くなる、という議論がありうる。しかしこれは日本の多施設少数患者試験には全く当てはまらない（大橋, 1998）。1 施設内で同様の適格条件を有した試験が並行に行われることから患者選択のバイアスはいるからである。実薬対照および主治医評価は、一見試験を pragmatic な立場から行う方法であるかのように思われる。しかしこれまでの日本の臨床試験の第 III 相試験の組み立ては、適格条件設定・評価・解析とも短期・小中規模な explanatory な試験のそれである。この中に、感度と信頼性が保証されないまま 2, 3 を取り込んだことに問題があったと考える。一方、(QOL を考慮した) 生存や疾患発症・再発をエンドポイントとした大規模試験を pragmatic に行うインフラストラクチャはわが国には未だに不十分である（大橋, 2003）。製薬会社主導で（現時点でのわが国の）GCP 準拠で大規模な pragmatic 試験を行うことはコスト上極めて難しい。そこまで厳しい質を求める必要もない。また併用も含んだ治療法選択といった目標からは、1 社スポンサーによる臨床試験には無理があり conflict of interests の観点からも望ましくない。一方、研究者主導でこれを行うには、研究費獲得・研究者のインセンティブ・施設の体制・データセンターおよびデータマネージャなど支援組織と人材の点でまだまだ問題が多い。市販後の研究者主導研究の体制強化とともに、研究の中立性を保ちつつ財政上の支援を可能とするシステムが必要であると考えられる。

## 3. なぜ疫学・臨床試験研究が日本で育たなかったか？

前書きにも述べたように、臨床医学系学会では Evidence Based Medicine (略して EBM) という言葉が 1990 年台に大流行した。これまでの医療科学に対する新しいパラダイムであるという熱狂から、反感・誤解にいたるまでさまざまなレベルの理解と態度が存在したが、現在では医療科学の重要な視点として一定の評価を得たものと位置づけられる。EBM とは、目の前の患者の問題点を一定の手順で定型化し、主に文献検索と抽出された文献の批判的吟味により過去の「証拠・根拠」を点検し、そこから有効な情報を引き出し、目の前の患者に対して実践することであ

る、とされている。そして、もっとも強力な evidence を提供するのが臨床試験とくに長期大規模試験そして複数の臨床試験を統計的に併合するメタアナリシスである。

我が国の臨床医学は、この evidence 作りにどれだけ貢献してきたか？

循環器系疾患の薬物療法に関して大きな貢献を果たした 288 の大規模臨床試験研究をコンパクトにまとめた「循環器メガトライアル The State of the Heart」(エルゼビア)の編者、松崎益徳教授(山口大学)は、引用した日本からの論文 0 という事態に対し、「本邦から報告されたものが皆無であるのは実に残念であり」と述べておられる。これは実は循環器に限らない実態であり、乳癌術後補助療法のメタアナリシス Early Breast Cancer Trialists' Collaborative Group (EBCTCG) の 133 試験中 5 試験、症例数 75000 例中 6239 例(1992 年当初)という貢献は極めて珍しい例である(EBCTCG, 1992)。何がそうさせたか？

「臨床研究を行う上で我が国にないもの」という問いかけに対し、当時の西條長宏薬物療法部長(国立がんセンター中央病院)はこう答えられている(1999 年 The 1<sup>st</sup> US-Japan Workshop on Clinical Trials の講演から)。「Everything!」。項目を挙げれば具体的には以下になるろう。

- (1) 医師研究者を支援するコーディネータ(CRC/リサーチナース)
- (2) プロトコルを書くことのできる医師研究者あるいはそのための教育
- (3) 生物統計学者(試験統計家)あるいは生物統計学の教育
- (4) 効率的なデータマネージメント・システムと中立的なデータセンター
- (5) Peer review あるいは査察によるデータの品質保証体制
- (6) 医師研究者主導型の臨床研究に対する法的規制
- (7) Medical writing の人材と教育
- (8) 臨床試験の意義の理解と患者さんの参加に対するインセンティブ
- (9) 臨床試験に参加する医師研究者へのインセンティブ
- (10) 研究資金とくに公的研究資金の供給と評価

さらに下部構造の観点から整理を行えば表 2 によるになるろう。制度に具現化される研究に対する国家の態度、教育環境そして国民の意識というように、臨床試験のインフラストラクチャに関わる問題は根が深い。国際化の中で抜本的な変革が必要とされており、ようやくわが国でも文部科学省・厚生労働省・経済産業省一体となった治験・臨床試験振興策が具体化されようとしている(たとえば、新医薬品産業ビジョン, 2007)。この中では生物統計家の育成と研究への参画の必要性が強くうたわれている。前書きに挙げたようなようやく芽生えた教育体制を、臨床試験に関わる他職種の教育と有機的に結合させいかに成長させるかが、学会あるいは生物統計専門家に与えられた今後の課題である。

#### 4. 方法論の展開：ランダム化を題材として

##### 4.1 必ずしも理解されていないランダム化

臨床試験に関わる生物統計学の方法論的発展はこの 10-15 年著しい。ここでは例として最大の貢献とされるランダム化に焦点を当て、発展を概観したい。

臨床試験におけるランダム化とは、治療法選択において恣意的な医師の選択あるいは患者希望

## 表 2. なぜ疫学・臨床試験研究が日本では育たなかったか？

- ◆学問としての認識の低さ
- ◆システム
  - 研究費配分と評価（厚生省と NIH の違い）
  - 支援スタッフ（CRC, 統計家, データマネージャ）
  - 国民の認識・研究参加のインセンティブ・ボランティア意識
  - 新薬許可制度
- ◆教育 70 年代の疫学・生物統計学教育へのアメリカの投資
- ◆風土 「データでものをいう」か、「なせばなる」か

を避け、複数の治療介入のいずれかに確率的な要素を伴って患者を割り付ける操作である。これによって、アウトカムに影響する患者背景や病態（がんの場合ならがんの性状）が治療群間で偏ることから生ずるバイアスを減らし、あるいはバイアスを平均的に除去し、正確な治療法間の比較がなされるとされる。しかし、その数学的な意味が十分理解されているとは限らず、またその普及も必ずしも順調ではなかった。このことは、医療関係者の中で未だにランダム抽出との区別がしばしばなされていないこと、臨床試験の教科書でさえも統計的推測の基盤としてランダム抽出を説明していることが多いことに見取れる。下記の Senn (1997, Glossary) の強烈な批判からも、上記の混同・不徹底が広く存在することが窺えよう：ランダム抽出は…多くの統計理論において理論上重要な概念であるが、臨床試験においてはほとんど現実味をもたない、私見であるが、不用で誤解を招く概念。

ランダム化が初めて提唱されたのは、1930 年代にイギリスの R. A. Fisher が、新種や肥料を評価する農事試験を対象として創案した「実験計画法」においてである。この実験計画法は 1947 年に英国 Medical Research Council によって開始された結核患者に対するストレプトマイシンの評価に採用され、その成功を通じ、臨床試験においても方法論としての有効性が確立されたとされている。しかし、1980 年代からランダム化試験の重要性を訴えるわれわれ統計家に対して、日本の臨床家から投げかけられた言葉は以下のものであった。

80 年代 「がんでランダム化試験ができるか」

90 年代 「外科でランダム化試験ができるか」

今日のわが国では、胃癌における脾摘の有無、大腸癌における開腹・内視鏡手術の比較など、がんの分野では外科手術にいたるまでランダム化試験が普及しエビデンス発信に貢献を果たしている。隔世の感がある。現在の臨床家の疑問は

00 年代 「しばしば小規模で患者を限定したランダム化試験の結果が、あるいは海外のランダム化試験の成績が、目の前の患者に適用できるのか」

という一般化可能性の疑問に進化をとげている。実は欧米でもランダム化試験の定着はそれほど

Jpn J Biomet Vol. 28, Special Issue 1, 2007

古いものではない。New England J. Medicine に錚々たる統計家が連名で発表した 1976 年の特別レポート (Byar et al., 1976) において、(歴史対照の活用や Cox 回帰などの背景因子調整の統計解析手法が現れても) ランダム化はもっとも信頼できる評価法として存続する、という宣言がなされているほどである。

ランダム化の驚くべき特長は、現在の用語を用いれば反事実的 (counterfactual) 想定の下で「実際は治療を行わない治療効果も (平均的には、正確に) 推定できる」という点である。また、ランダム化によって初めて検定の  $p$  値が計算可能になる、という数学的構造は医学関係者には殆ど理解されていないように思われる。

#### 4.2 ランダム化に関連する研究成果とその応用

筆者自身、臨床試験方法論に関する研究はランダム化に代表される患者割り付けに関するものが中心であった。筆者と共同研究者との、これら研究成果をざっとあげてみると以下ようになる：

- (1) がん臨床試験において提唱された、患者背景の群間分布を均衡させる動的割り付け法の登録センターによる実施と他臨床分野への応用 (わが国では初めての大学の登録センターにおける実施)、および理論的に曖昧であったその解析方法に対する提唱
- (2) がん臨床試験において提唱された Continual Reassessment Method の改良と応用、とくに臨床家の判断を容易とするための図的表現の工夫と、新薬申請を早めるための臨床試験第 I (第 I/II) 相試験への実際の活用
- (3) Web を活用した臨床試験登録と割り付けの世界初めての実施
- (4) コンピュータネットワークを用いた薬剤管理と割り付けの実施
- (5) 施設と時期組み合わせによるクラスターランダム化 (世界最初の試行例と思われる)、可変ブロックと患者 1 例毎の鍵との組み合わせ

がん臨床試験においては、患者背景が患者予後に強い影響を有するため、また循環器系試験などと比較すると比較的小規模な試験が多いこと、中間解析を行う必要があることから、試験のどの時点でも治療群間で患者背景をバランスさせる動的割り付け法が欧米においては 1980 年台中ごろから応用されてきた。筆者らはこの方法論を実応用例とともにわが国に紹介する (坂本他, 1988) とともに、この実施を行うための割り付けセンターを大学教室内に開設し、ノウハウを蓄積するとともに、治験・市販後臨床試験・研究者主導試験の実施に協力してきた (現在この機能は NPO 日本臨床研究支援ユニットに移行されている)。このノウハウは、現在では国立がんセンター・国立大学病院医療情報ネットワーク (UMIN)・民間の CRO 等で広く活用されている。またこのデザインから得られたデータを正確に解析する方法論として、並べ替え検定が適用可能であることを指摘し、このためのコンピュータソフトウェアを開発した (Ohashi, 1990)。現在では一般的であるが、Web 活用による患者登録と動的割り付けは、文献としては世界で初めて Kiuchi et al. (1996) によって発表された。国立大学病院医療情報ネットワーク (UMIN) 上で実装され、現在でも広く活用されている。1990 年中ごろまでのわが国には、(がん術後補助療法を除けば) 対象数 1000 例、期間 3 年を超える長期臨床試験の経験とくに治験はほぼ皆無であった。長期臨床試験において 2 重盲検維持と有効期限遵守を行いつつ薬剤管理を行うために、わが国では初めて、登録センター (東京大学医学部疫学・生物統計学教室)・製薬会社工場・製薬会社開発部をネット

ワークで連結するシステムを構築し、実際に運用した例が Gotoh et al. (2000) である。

とくに救急患者を対象とした臨床試験においては、丁寧な説明の後に個別の同意を得ること、治療開始を遅らせることなく患者個別にランダム化を行うこと、割り付けられた治療を正しく実施することがしばしば困難である。これを解決するために、各施設において月を単位としてランダム化を（中央の登録センターからの指示により）行い、その月に登録された患者に対する同意説明は簡略化し、すべて同様の治療を行うランダム化の提案が (5) 前半である。このデータ解析には工夫が必要であり、また恣意的な患者選択が行えないよう監査等の仕組みが必要であるが、実際にある循環器系の試験において成功裡に活用できた。また、小児対象のワクチンの 2 重盲検開発試験において、すべての結果が出揃うまで結果を開示しないと対象児のワクチン追加が遅れるため、当該児のデータ固定がすみ次第、一例毎に開鍵するデザインを提唱し、成功裡に活用できた。これが (5) 後半である。このためには割り付け結果の予見性を防ぐ工夫が必要であるが、ブロックの大きさを確率的に変化させる可変ブロック法の採用により解決が可能となった。これらの成果については現在論文に取りまとめ中であるが、臨床試験への要求水準が高くなるにつれ、臨床側のニーズに応える新しい統計的手法が必要とされ、さらにこの応用により臨床試験から提供される evidence 水準が高くなる、という螺旋的展開がわが国でも起きつつあるように実感する。この加速のためにも、臨床側のニーズを適格に把握できる試験統計家の育成が必要である。

## 5. 臨床試験のパラダイムシフト

他の応用統計の分野と異なる臨床医学統計の究極の課題、あるいは醍醐味は、一般化と個別化の「行きつ戻りつ」であろう。population pharmacokinetics に代表される pharmacokinetics の統計解析の分野はこの良い例である（詳しくは大橋, 2005）。

次の挑戦課題は薬物の作用点・受容体以降の個体差に対してである。治療法評価のための標準手順であるランダム化とは、患者の個体差は認めた上で公平な (valid) 判定の場を人工的に構築する実験技法であり、これにより集団としての、あるいは平均としての治療法の評価が可能となる。この結果を個別の患者にどう適用するかは医療のアートの部分とみなされてきたが、これを問題解決技法として体系化しようとしたのが EBM であろう。ランダム化試験の対象は、とくに同意という選別プロセスを経るがゆえに限定された特殊な患者集団であり、この結果がどれだけ素直に（躊躇無く）患者一般あるいは目の前の患者に適用されるかは、統計学においては一般化可能性（外的妥当性, generalizability）の問題と呼ばれてきた。表 3 に一般化可能性を検討する、あるいはこれを高める方法をまとめた。メタアナリシスあるいは Richard Peto に代表されるイギリスの統計家が simple and large scale evidence と呼ぶ適格条件をゆるくした大規模試験は、平均的な効果に対する結論の頑健性を高めようとする方向であり、個別化治療とは、逆にサブグループ解析等により治療の予測因子を抽出し、これにより治療方法は個別化するものの全体として治療効果を上げる、あるいは無駄な副作用を避けようとする方向である。

以下に、臨床試験の個別化医療へのパラダイムシフトを実感する体験を紹介したい。先にもあげた乳癌術後補助療法メタアナリシス EBCTCG の会議の顛末である。

EBCTCG は、Richard Peto の主導のもと、世界中の研究者から乳癌補助療法のランダム化試験

表 3. ランダム化試験結果の一般化可能性を高めるために

- ◆患者背景の解析  
結果の差異の説明
- ◆部分集団の解析  
部分集団における効果の差、交互作用の検討
- ◆再試験  
FDA の方針
- ◆メタアナリシス  
複数の独立な研究結果の統計的併合
- ◆緩い選択条件  
large-scale-randomized-evidence
- ◆ITT (Intention-to-treat) 解析

の個票データを集め、これをメタアナリシスの手法により解析し、1992年のLancetの論文以来、1995年、2000年と5年毎にデータ更新・解析・出版を行い、乳癌術後補助療法の重要エビデンスを提供してきた。ところが2005年の会議の直前(2005年7月5日)に、オックスフォードから全参加予定者にショッキングなメールが送られた。EBCTCGが存亡の事態に立ち入っているというのである。2000年の解析結果がようやく2005年に発表されるという出版遅れが生じ、また、データ提供を行ってきた研究グループの要求に沿った解析に、オックスフォードの統計グループなかなか応じようとしてこなかった。そこで不満が爆発した。世界を代表する研究グループ(BIGのM. Piccart, CALGBのE. Winer, ECOGのG. Sledge, IBCSGのA. Coates, A. Goldhirsch, R. Gelber, NCI-CanadaのK. Pritchard, NCCTGのE. Perez, NSABPのE. Mamounis, SWOGのK. Albain, TBCIのW. Wood)が、連名で「このままならデータ提供を行わない」という最後通牒を突きつけたのである。彼らはこれまでのEBCTCGの功績は認めるものの、以下のような主張を行った：

- メタアナリシスを必要とした昔と異なり、いまや試験は大型化し、単独で答えを出せるようになってきている。
- そして集団全体に対する最適な治療というより、(分子マーカーによって定まる)生物学的に意味のあるサブ集団に対して最適な治療法を決定する時代にきている
- trastuzumab(ハーセプチン)や、2005年のASCO(アメリカ臨床腫瘍学会)で進行・再発乳癌に延命効果が証明されたbevacizumab(アヴァスチン)など、今や有効な治療法が続々登場している。これまではオックスフォードに頼ってきたが、自分達でメタアナリシスを行い重要な解答が得られるようになってきた

一騒ぎの後、データ提供者(会議参加者)へのアンケートに基づき、臨床家・統計家からなる国際的Steering Committeeが結成され、オックスフォード統計家グループと協力・連携しながらメタアナリシスが進められることとなった。2005年の会議ではデータ提供の遅れにより不十分な会議となってしまったものの、2006年6月のSteering Committee(座長はK. PritchardとM. Jpn J Biomet Vol. 28, Special Issue 1, 2007



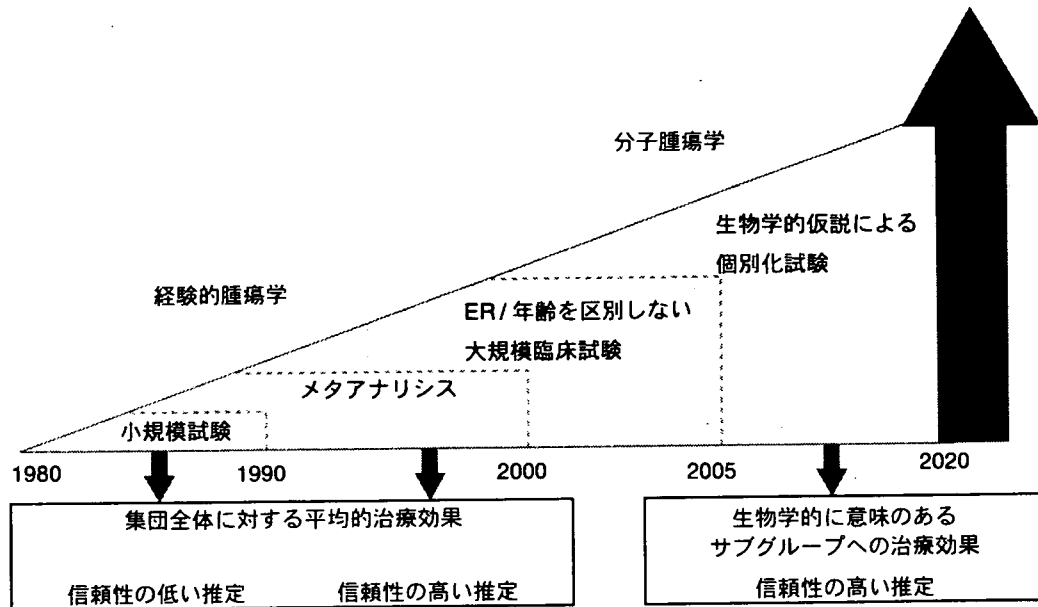


図 1.

Piccart, 日本からは、オックスフォードの指名により著者が参加) を経てデータは再び着実に収集されるようになり、2006年9月11日から体制を一新しての会議が無事開催されることになった。具体的には、Petoをはじめとするオックスフォード統計家が一方的にプレゼンテーションするこれまでの形式に変わり、臨床家が解析結果を受けて解釈のコメントを述べる形となった。

図1に、上記の議論の背景となった乳癌補助療法臨床試験の歴史的展開とパラダイムシフトを挙げた。1980-90年代には統計的な詰めが甘く症例数の小さな試験も多数行われ、集団全体に対する平均的な治療効果を精度高く推定するためメタアナリシスが必要とされた。1990年代から2000年代には、統計的にきちんと設計され症例数は多いものの年齢やERに制約を置かない大規模試験が数多く行われた。そして trastuzumab など分子標的薬の登場、DNA マイクロアレイに代表される分子生物学の知見の活用の時代となり、生物学的に意味のある部分集団を標的とした臨床試験が設計される個別化治療の時代となった。M. Piccart 率いる BIG では、マイクロアレイ解析で同定される高グレード群に限定して、G-CSF を用いて2週毎の投与を行う dose-dense 化学療法と通常の化学療法 (AC-Taxol) を比較する大規模臨床試験を開始している。効果がホルモン依存性による可能性が高いため、全体である程度の効果があるのか、ER-に限定した場合にのみ強い効果が発現するのかを、前向きに「統計的な多重性」を考慮し Hochberg 法を用い評価するデザインとなっている。このような標的を限定した臨床試験結果も将来はメタアナリシスの対象となり、さらに細かくサブグループ解析されることにより、新たな仮説を生み出すことになる。

## 6. 終わりに

がんの分野では Imatinib や trastuzumab・bevacizumab など、-nib (小分子の分子標的酵素阻害薬) や -mab (モノクローナル抗体) の開発成功を受け、ベンチャー・メガファーマ合わせて抗が

Jpn J Biomet Vol. 28, Special Issue 1, 2007

表 4. 最近のがん臨床試験デザイン上のアイデア

- ◆ベイズ流手法の導入
  - Data-driven とすることで柔軟性を確保
  - CRM, 第 II 相試験への応用, 第 II 相から第 III 相へのスムーズな移行  
(Lee and Feng *JCO* 2005;23:4450-7 の文献 32-36)
  - 日本でも CRM による I/II 相試験で承認例 (Bortezomib)
- ◆第 II 相試験とくにランダム化 II 相試験の工夫
  - Lee and Feng *JCO* 2005;23:4450-7
  - スクリーニング第 II 相試験 Rubinstein et al. *JCO* 2005;23:7199-206
- ◆感度を高めるための Enrichment design
  - ランダム化中止実験
  - Freidlin and Simon *JCO* 2005;23:5094-8
- ◆標的集団を意圖した検証
  - Simon and Maitounam *Clin Cancer Res* 2004;10:6759-63
  - 日本でも研究者主導試験から治験へ応用がはじまりつつある

ん剤の開発が加速されている。現在アメリカで開発されている新薬候補のおよそ 2/3 が抗がん剤であるという。分子生物学的知見を用いた標的探索とこれに引き続く high-throughput スクリーニング技術の活用、上記のベンチャー参入により開発候補物質そして有望な薬剤組み合わせの数は爆発的に増加している。また従来の cytotoxic 薬（細胞毒薬）を対象として組み立てられた臨床試験のストラテジーが見直され、マイクロアレイに代表されるマーカー活用とこれによる個別化医療も考慮した表 4 のような新しい臨床試験方法論が数多く提案され、活用されようとしている（詳しくは大橋, 2006）。新 GCP による国際標準の導入に次ぐ大きな変革がわが国の臨床試験の世界では起きつつある。本稿では詳述しなかったものの、これまでわが国では実現困難とされた循環器系大規模（予防）臨床試験の成果も 2006 年から着々と発信されつつある。これまで全く実施されてこなかった検診のランダム化臨床試験も乳癌を対象として開始された。繰り返し主張したように、研究を「螺旋的に展開させる」生物統計専門家あるいは試験統計家がますます広い分野で要求されつつある。

#### 参 考 文 献

- Byar, D., Simon, R. M., Friedewald, W. T., Schlesselman, J. J., DeMets, D. L., Ellenberg, J. H., Gail, M. H., and Ware, J. H. (1976). Randomized clinical trials: Perspectives on some recent ideas. *New England Journal of Medicine* **295**, 74-80.
- Early Breast Cancer Trialists' Collaborative Group (EBCTCG). (1992). Systemic treatment of early breast cancer by hormone, cytotoxic or immune therapy. *Lancet* **339**, 1-15.
- Gotoh, F., Ohashi, Y., and the Cilostazol Stroke Prevention Study group. Design and organization of the Cilostazol Stroke Prevention Study. (2000). *Journal of Stroke and Cerebrovascular Diseases* **9**, 36-44.

- ICH E9 Expert Working Group (1999). Statistical Principles for clinical trials. *Statistics in Medicine* **18**, 1905-1942.
- Kiuchi, T., Konishi, M., Bandai, Y., Kosuge, T., and Ohashi, Y. (1996). A World Wide Web-based user interface for a data management system for use in multi-institutional clinical trials-Development and experimental operation of an automated patient registration and random allocation system. *Controlled Clinical Trials* **17**, 476-493.
- 厚生労働省 薬審第 1047 号. 臨床試験のための統計的原則 (E9). 1998 年 11 月 30 日.
- 厚生労働省. 新医薬品産業ビジョン. 2007 年 8 月 30 日.
- Ohashi, Y. (1990). Randomization in cancer clinical trials : Permutation test and development of a computer program. *Environmental Health Perspectives* **87**, 13-17.
- 大橋靖雄. (1998). 有用性と超多施設試験の意義. *統計数理* **46**, 122-127.
- 大橋靖雄. (2003). 「市販後臨床試験」と「市販後の臨床試験」. *薬剤疫学* **8**, 45-50.
- 大橋靖雄. (2004). クリニカルトライアル-2020 年への展望. *遙か* 創刊号, 38-45.
- 大橋靖雄. (2005). 個別化医療 : 統計的視点. *Molecular Medicine* **42**(8), 912-917.
- 大橋靖雄. (2006). 個別化医療 : 癌臨床試験デザインの最近のデザイナー分子標的薬剤を中心として. *Mebio Oncology* **3**(3), 62-71.
- Piantadosi S. (1997). *Clinical Trials: A Methodologic Perspective*. Wiley.
- Pocock, S. (1989). *クリニカル・トライアル (コントローラ委員会訳)*. 篠原出版.
- Redmond, C and Colton, T(ed.). (2001). *Biostatistics in Clinical Trials*, Wiley.
- 坂本純一, 中里博昭, 大橋靖雄. (1988). 最小化法 (Minimization Method) の癌比較対照試験への適用. *癌と化学療法* **15**(6), 1935-1942.
- Schwartz, D., Flamant, R., and Lellouch J. (1980). *Clinical Trials* [translated by Healy MJR]. Academic Press.
- Senn, S. (1997). *Statistical Issues in Drug Development*. Wiley.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison Wesley.

〈抄録〉第27回日本臨床薬理学会年会 2006年11月29日～12月1日 東京  
シンポジウム1：わが国の心血管病大規模臨床試験の現状と課題

## 1. 大規模臨床試験の品質管理 —MEGA Study を例として—

大橋 靖 雄\*

### 1. 臨床試験データの品質管理と品質保証

韓国のES細胞実験そしてわが国でも実験データの改竄・捏造が話題となり、研究結果の信頼性をいかに保証するかが議論されている。誤解を恐れず大胆に言えば、基礎研究結果の信頼性・信憑性は追試によって最終的には確認される。一方、大規模な臨床試験や疫学研究においては追試が極めて困難なため、その研究自体で信頼性を確保すること、さらにその信頼性の高さを他人に認めてもらう必要が生じる。臨床試験の品質管理・品質保証において、ISO流の「プロセスあるいはシステムとしての保証」の考え方が導入された背景がここにある。

### 2. MEGA Study 実施のためのシステム

日本人を対象として、軽中度の高コレステロール血症をターゲットとし、プラバスタチンを用いてCHD発生抑制をプライマリエンドポイントとして計画・実施された1次予防試験がMEGA Study<sup>1)</sup>である。これは1993～94年に厚生労働省の研究事業として開始され1995年からは三共株式会社が受け継いで実施された。最初の患者登録は1994年2月、最後の登録は1999年3月、全例がほぼ5年追跡される2004年3月をもって追跡終了とされた。

試験の実施は困難を極めた。1995～96年まではわが国の臨床試験実施基盤は脆弱であり、登録も難航した。データセンターと試験事務局の整備強化とくに施設へのコーディネータの派遣、(データに直接関わらないことは一貫しつつの)三共の支援体制の強化、啓蒙書や栄養パンフレット配布・食事会の開催などの患者に対するインセンティブ、新GCP導入に伴う試験環境の変化等により試験途中から登録数は飛躍的に上昇し、当初不可能と思われた目標8,000例の登録数が達成された。

試験の中立性と品質保証を行うため、本試験では以下のようなシステムを採用した。

- ・スポンサーから独立した試験組織、とくに研究事務局とデータセンターの設置
- ・市販後試験では比較的珍しかったモニタ派遣とコー

ディネータ派遣

- ・試験結果をモニタし、中止やプロトコル変更を勧告する独立モニタリング委員会の設置
- ・米国CDC認証によるコレステロール値(総コレステロールとHDL)の中央測定
- ・厳正な中央割付け
- ・濃密なデータマネージメント
- ・参加患者の同意を円滑に進めコンプライアンスを高めるためのインセンティブ
- ・患者適格性判定のための評価委員会(27回開催)
- ・盲検化でのイベント評価委員会(6回開催と持ち回り判定)
- ・消息不明例に対する住民基本台帳閲覧による転居先・生死確認

最後の住民基本台帳閲覧を除けば、いずれも今日では大規模臨床試験の常道ではあるが、当時これらを大規模に採用した点でMEGA Studyは画期的であった。また徹底した追跡により、(同意撤回による追跡終了を追跡完了と換算すれば)追跡不能例はわずか102例1.3%であった。以下、若干補足しよう。なお、コーディネータ(CRC)やモニタ(CRAとよばれた)の業務、データセンターと事務局の業務、各種委員会活動などMEGA Study実施の詳細については、Progress in Medicine増刊号「完成したMEGA Studyのすべて」2006; Vol. 26 suppl. 2に詳しい。

#### 1) 研究組織

臨床試験の実施組織は、大きく

- ・大規模な試験ごとに研究組織を構築する循環器系型
- ・小規模な試験を数多く実施するのに適した、恒常的な事務局・データセンターを擁すがん共同研究型に分けられる<sup>2,3)</sup>。ともに米国のNIHの支援による多施設共同研究を実施するために1970～1980年代に確立したシステムである。これらは1986～88年頃に筆者らによって日本に導入された。MEGA Studyは本格的に前者のシステムを採用した初期の例である。以下がMEGA Studyの特徴である(Fig.参照)：
- ・通常はデータセンターの下におかれるイベント判定委員会、消息調査委員会、評価委員会をわが国のこれまでの治験の習慣になじむよう運営委員会の直轄としたこと

\* 東京大学大学院医学系研究科生物統計学/疫学・予防保健学  
〒113-0033 東京都文京区本郷 7-3-1

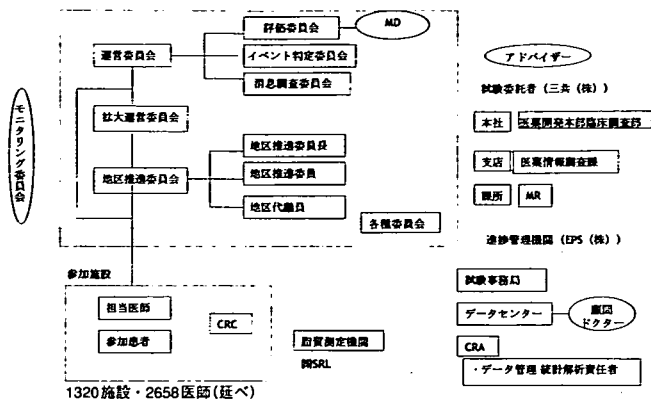


Fig. MEGA の試験組織

- ・試験実施を支援する、しかしデータにはいっさい触れないスポンサー組織側の存在
- ・事務局内にモニターと臨床試験コーディネータ・栄養指導を行う栄養士を配したこと

評価委員会は患者の適格性判定と統計解析上のデータの扱い（採用あるいは除外）をおもに決定するために設置された委員会である。これも、試験総括医師（PI；Principal Investigator）を中心とした少数の医師グループとスポンサー側とが協議のうえでデータの扱いを決定してきたわが国の治験の習慣に基づいたものである。データセンターの経験がまだ不十分でデータの取扱いを決定しながら試験を実施せざるをえなかった状況下では、このような組織は不可欠であった。

なお、わが国で医療施設へのコーディネータの配置が本格的に始まったのは新GCP完全施行の1998年以降であり、しかも市販後の試験であることからMEGA Studyでは施設に所属するコーディネータを活用することは不可能であった。またコーディネータを派遣するSMOも未成熟であったことから、上記のように事務局にコーディネータを配し施設に派遣する形となった。患者同意を得る上でまたデータ管理を行ううえで、コーディネータの活躍は注目すべきものであった。

## 2) データマネジメント

ようやく最近、厚生労働省も臨床試験を支えるデータマネジメントの重要性を認識し、治験あるいは臨床試験基盤整備の要件に含めるに至ったものの、この言葉がわが国に筆者により導入された1985～86年から現在までの経過は順調ではなかった。専門性の認知はなかなか進まず、教育体制も不十分、専門書もわが国で出版されたものはわずか1件<sup>4)</sup>という状況であった。

MEGA Studyでは、CRO (EPS社) がMEGA Study 専用に設立したデータセンターにおいて、標準的なデータマネジメント・プロセスにほぼ準拠し、紙のCRFを用いたデータマネジメントが行われた。試験期間中のCRFの枚数は1例平均12.5枚で

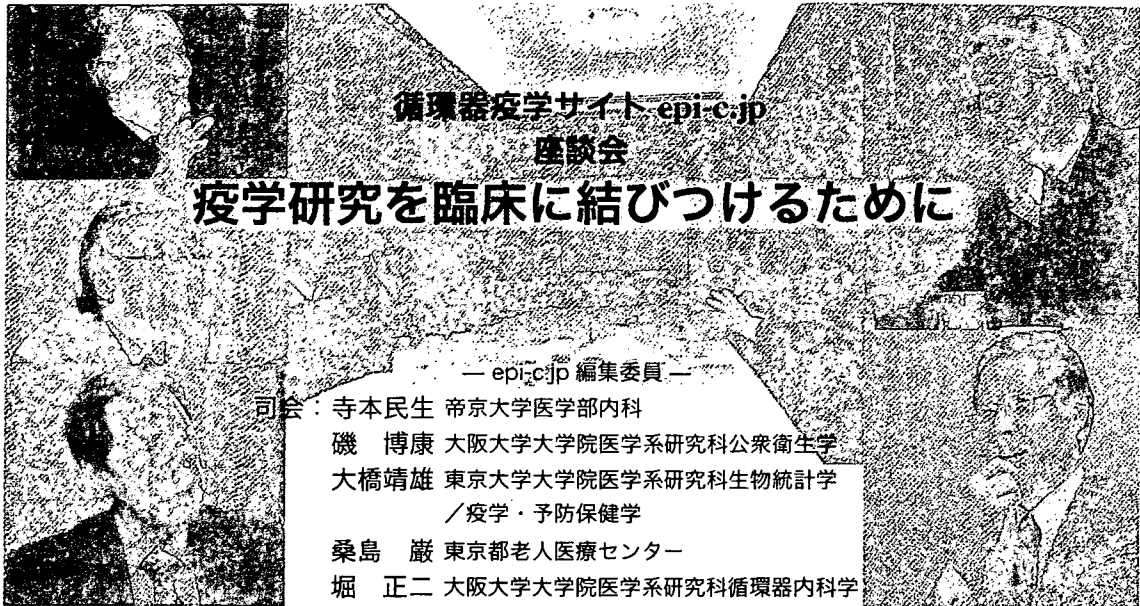
あったが、観察期CRF、イベント調査票・心電図連絡書・転帰連絡書など付随の書式を併せると1例当たり40～50枚の書類が発生し、さらに入力・問合せのために発生する書式を併せ、書類の総重量は45tにも達した。おそらく今後これだけの大規模なデータマネジメントを紙ベースで行うことは不可能であろう。大規模となった理由としては、試験全体が長期大規模であったことに加え、施設の体制が不備であったことによる負担増が大きい。データマネジメントを効率化する最大の方策は、収集するデータを絞ると同時に、発生源すなわち臨床現場でのローカルデータマネジメントの水準を上げ、一方CRF設計・データ処理システムの構築・問合せ等のセントラルデータマネジメントの質も同様に上げ、両者の連携を密とし、問合せと書類の行き来を最小限あるいは不要とすることである。現在盛んに導入が試みられているEDC（電子媒体への直接入力）も効率化の方策ではあるが、ローカル・セントラル両者の効率化・標準化そして質向上が前提であり、これ抜きの導入は施設側の負担増につながるのみである。

## 3. 大規模臨床試験を成功させるために

大規模臨床試験がわが国で高品質で実施可能であり、この成果が国際的に発信できることを示した点でMEGA Studyの意義は大きい。盛んになりつつある研究者主導臨床試験を支えるために、試験の計画を担う医師とともに、データマネージャに代表される支援スタッフの教育を本格的に考えるべき時期に来ていると筆者は考える。データマネージャに関しては日本科学技術連盟では毎年3月に初心者向け短期コースを開催しているものの、これだけでは不十分である。これも毎年開催されているDIA (Drug Information Association) のデータマネジメントのフォーラムは今のところ企業の治験担当者向けである。ようやく横浜市立大学で臨床試験支援スタッフ (CRC, モニター, データマネージャ) 育成のための専門職大学院設立を構想中であり、東京大学では臨床家に対する1年教育も行う専門職大学院・公共健康医学を2007年4月に開講した。また臨床試験データ管理学寄附講座も活動を開始した。すでに東海大学と京都大学には、医師を対象とした臨床研究の教育プログラムが成果をあげている。個人的にはこれらの連携による医師・支援スタッフ双方の教育システムが有機的に機能することを願っている。

## 文 献

- 1) Nakamura, et al. *The Lancet* 2006; 368: 1155-63.
- 2) 大橋靖雄. 医学のあゆみ 1988; 146: 11-3.
- 3) 大橋靖雄, 荒川義弘(監修). 臨床試験の進め方. 南江堂, 2006.
- 4) 大橋靖雄(監修), 辻井敦(著). 臨床試験データマネジメント. 医学書院, 2004.



循環器疫学サイト epi-c.jp

座談会

## 疫学研究を臨床に結びつけるために

— epi-c.jp 編集委員 —

司会：寺本民生 帝京大学医学部内科  
磯 博康 大阪大学大学院医学系研究科公衆衛生学  
大橋靖雄 東京大学大学院医学系研究科生物統計学  
／疫学・予防保健学  
桑島 巖 東京都老人医療センター  
堀 正二 大阪大学大学院医学系研究科循環器内科学

1990年代に確立されたEBMという手法は、その後の夥しい臨床試験データとともにわが国に導入され、医療現場に多大な影響を与えてきた。わが国においても、日本人のエビデンス蓄積の必要性が認識され、医師主導の臨床試験も数多く実施されるようになった。それらの動きに伴い、対象母集団の基本情報が介入試験実施の前提として重要視されるようになった。疫学観察研究の成果は、いまやEBMを支える両輪の一つとして、臨床試験の評価と適応に欠くことのできない位置を占めつつある。ここでは「循環器疫学サイト epi-c.jp」の編集委員により、疫学研究が臨床に果たす役割と可能性について討議してもらった。

**寺本** 疫学研究はなかなかはじめなかったのですが、大規模臨床試験の結果が臨床の場でどれだけ生かせるかを考えるためには、疫学の知識によって解釈する必要性を感じるようになりました。今日は疫学研究を臨床医はどうみていったらよいのかを話していければと思います。

最初に磯先生に循環器疫学の現状についておまかな紹介をしていただきたいと思います。

**磯** 循環器疫学では、フラミンガム研究とSeven Countries Studyがパイオニア的な研究と位置付けられます。日本でも久山町研究、秋田県の井川町研究や大阪の勤労者コホート研究が昭和30年代後半からはじまっています。フラミンガム研究からは危険因子という概念が、久山町研究や井川町研究などからは日本の脳卒中の実態が示されました。

一方、癌の疫学あるいは生活習慣病の疫学研

究として、もう少し大がかりな疫学研究が20年ぐらい前から世界各国ではじまりました。米国では10万人規模の看護師のコホート研究であるNurses' Health Studyや、男性の歯科医師、獣医など5万人規模のコホート研究であるHealth Professionals Studyなどです。

それまではフラミンガム研究など循環器疾患の研究は数千人からせいぜい1万人ぐらいの規模でしたが、栄養や運動などの生活習慣を危険因子として評価するために大規模になってきました。日本では国立がんセンターと国立循環器病センターに大学などの研究機関が加わったJapan Public Health Center-Based (JPHC) Prospective Studyという生活習慣病のコホート研究があります。また名古屋大学が中心になって全国の衛生・公衆衛生の研究者が行っているJapan Collaborative Cohort Study for Evaluation



磯 博泰 氏

for Cancer Risk (JACC) も 10 万人規模の大きな疫学研究です。茨城県でも 10 年前から約 10 万人規模でコホート研究が行われています。

最近では従来の数千人規模のコホート研究と、10 万人規模のコホート研究をメタアナリシスする研究として、大橋先生が事務局をされている Japan Arteriosclerosis Longitudinal Study (JALS) 研究があります。

**寺本** いま磯先生からいくつかの疫学研究が紹介されましたが、桑島先生は臨床の側から疫学をどのようにみていますか。

**桑島** 疫学というと、やはりフラミンガム研究がわれわれ循環器医にとってはすぐに思い浮かぶし、日常診療で非常に参考になっています。

フラミンガム研究で 40 年も前に提起された危険因子の重複 (clustering) という概念は、近年注目されているメタボリックシンドロームに結びついています。血圧に関しても、フラミンガム研究などの疫学研究から、血圧と心血管合併症の関連について知見が得られるようになりました。

しかし、臨床医はこれまでどちらかというところ疫学研究を軽視する傾向がありました。高血圧の分野では、すでに 20 年前に疫学研究により血圧は低い方が心血管合併症の発症が少ないことが報告されていたにもかかわらず、多くの臨床家は「治療」は別の問題、血圧はあまり下げなくてもよいと考えていたのです。しかし、最近になってようやく血圧は下げるべきだと考えられるようになってきました。

疫学的知識が治療に結びつかなかった理由の一つは、よい治療がなかったために血圧や脂質

値を下げるができなかったからです。いまはスタチンがでてきて、降圧薬もよいものがでてきたため、血圧や脂質を厳格に下げることによって心血管合併症を予防できることが証明されたのです。つまり疫学の正しさがやっとな治療学で証明されたということです。

一方、私は最近、ランダム化比較試験 (RCT) は限界にきているという印象をもっています。治療学が発展し、試験方法も PROBE 法なども行われるようになってきたため比較しても群間に差がつかなくなりました。いまこそ疫学の成果を踏まえた臨床研究の構築が大事なのだといえます。

### 「文化」としての疫学、「文化」を反映する疫学

**寺本** 堀先生、いかがですか。

**堀** 疫学には各国のバックグラウンドの違いができるように思います。スウェーデンやノルウェーでは、国勢調査のようなもののなかで健康データを一緒に集めていて、これが北欧の疫学のベースの 1 つになっています。北米ではフラミンガム研究がスタートした後、カナダのマクマスター大学で「疫学」を学問体系にする取り組みがはじまりました。

一方、日本はというと、いまだに「この病気が何人いるか」と聞いても数字が出てこない、きわめて特異な国です。自分の実態を知ることに関心をもってこなかったといえます。医療というのは自分の目の前に来た患者さんを治すことという意識が非常に強くて、関心が「個」だけに向かっていた。

疫学というのは文化だと思うのです。ライフスタイルが疫学のベースを作っているのです。どの病気が何人いるのかも知らないということは、自分の国にどんな文化があるのかを知らないということです。欧米と日本では実態が違います。日本はどうなのかということをもっと知らない、あらたな治療法が生まれてきても臨床の場で生かせません。このことに皆が気づきはじめているのではないかと思います。

**寺本** あるライフスタイルのなかで起こってくる病気というのがあるわけですから、日本のなかで起こってくるものと、欧米で起こってくることは違うわけですね。大橋先生、いかがですか。

**大橋** 大変刺激的なお話です。治療学が進んで疫学データが脚光を浴びるというのはまさにそのとおりです。

また、疫学に文化の違いがでるということには大賛成です。日本は農耕民族の文化で、周りと合わせることを考えてきました。それが狩猟民族とはまったく違いデータに基づいて行動するということはあまりなかった。私の先生である開原成允先生が「それがデータベースができなかった理由である、文化の違いだ」といわれていました。国の統計の利用度は、日本はOECD参加国のなかで最低です。

しかし、いま日本の疫学は大変おもしろい時代にはなっています。日本では健診データをベースに疫学研究を行っています。このような手法をとっている国は世界にあまりないと思うのです。疫学が大きく変わってきて、健康増進にもリンクしていく。ダイナミックな時代になってきています。

### リスクファクターと『ベネフィット』ファクター

**寺本** 癌や生活習慣病の疫学では住民健診データから行っているのですね。

**磯** 日本では自治体の住民健診をベースにして研究するという手法が多いですね。欧米では地域でランダムサンプリングし、電話でリクルートして個別的に健診するという手法がとられています。日本でもそのようにできないことはないとは思いますが、いままでの文化的な背景が違うために一般的には行われていません。集団健診を活用して研究するという手法が使われ、その伝統を踏襲しているといえます。

**寺本** 生活習慣の研究のなかで非常にインパクトがあるなと思ったのは、磯先生がお示しになった、魚の摂取と冠動脈疾患を調べた JPHC



堀 正二 氏

の結果です (*Circulation* 2006;113:195-202)。なぜ調べようと思われたのですか。

**磯** 魚や海獣類を多食しているグリーンランドエスキモーのコホート調査では心筋梗塞が少ないことが報告されていました (*Int J Epidemiol* 1988; 17: 514-19)。また Kromhout らがオランダ人の 20 年間のコホート研究により、魚を食べない人よりは週に 1 回か 2 回食べる人のほうが心筋梗塞は少ないと報告しています (*N Engl J Med* 1985; 312: 1205-9)。

日本人は魚をよく食べる民族で、米国と比べても心筋梗塞の死亡率は 4 分の 1 です。Seven Countries Study でみると、総コレステロール値は日本は非常に低いのですが、各国の間でコレステロール値が同じ場合でも心筋梗塞による死亡率は明らかに低い。コレステロール値だけでは心筋梗塞の国の差は説明できないと判断されます。

JPHC の結果では、週に 1, 2 回食べている人と毎日食べている人で比較したところ、毎日食べている人で、リスクが半減しているのです。欧米人では週 1, 2 回の摂取で心筋梗塞が半分になることがわかっていたのですが、毎日食べればさらに半分になるという可能性が出てきました。

**堀** 魚食に関しては“リスク”ファクターではなくて“ベネフィット”ファクター、よいファクターですよということを、日本からもっと発信しなければいけないと思うのです。日本は世界でもっとも長寿の国です。なぜ日本人が長寿なのか。その“ベネフィット”としてわかっていることは、リスクファクターの逆のものとし





寺本民生氏

て、世界に示すべきだと思うのです。

**寺本** 私もそう思います。日本の食生活はかなり変化が起こっていて、いまの子どもたちはあまり魚を食べなくなっている、つまりよい文化をいま失いつつあるのです。

**磯** あまり一般的ではありませんが、英語ではプロテクティブファクターという言葉があります。

**堀** もっと積極的な意味を持たせたいですね。

## RCTを疫学から読む

**寺本** 疫学、統計の先生に伺いたいもう1つの問題は、ランダム化比較試験（RCT）についてです。臨床研究の読み方は、先生方はどのように考えていらっしゃるのでしょうか。

**大橋** RCTでは一次評価項目をきめて、それが証明できるサンプルサイズを設定しています。つまり、一次評価項目以外に関しては確証とはなりません。データの読み方は疫学と同じで、試験が行われたのはどういう集団であるのかを考える。サンプルは選ばれているのですから、それが一般化できるのかどうか最大の問題です。

疫学には医療の問題点を指摘する役割もありますが、よい治療が適切に行われているかを評価するのも、疫学の重要な役割です。そのようなアウトカムリサーチにはデータを集約できるシステムが必要ですし、薬剤の製造販売後の薬剤疫学的な研究も必要です。つまり疫学は医療の最初と最後のところ、問題点の指摘と評価に役割があります。そのように使ってはじめて

RCTの結果をどう読むかにつながります。

RCTがどんどん難しくなっている点についてですが、サロゲートエンドポイントをどう評価するかというのは、疫学から考えなければなりません。そのような臨床試験と疫学との相互作用がありうるし必要であると思います。

**堀** RCTは臨床研究の手法として最終的なものでしょうか。それとも将来的には違う手法がでてくるのでしょうか。たとえば国民の登録がきちんとなされて、服用している薬、その時期、その背景情報も全部データベース化されたならば、それで治療法の評価研究はできますね。そうだとすると、RCTという、抽出された集団で行う研究手法はオールマイティではないということになりますね。

**大橋** おっしゃるとおりRCTはオールマイティではありません。しかし、RCTは妥当な（バイアスのない正確な）結果を常に出せるので大きな意義があります。それに代わるような、データベース構築ができればRCTはなくてもよいのですが、それは現在の知識水準では無理だろうと思います。特に薬剤の効果を見るにはRCTは残るでしょう。しかし、RCTでは選ばれた集団のなかでの効果は推定できるけれども一般化の保証は必ずしもない。この欠点は意外に理解されていません。その欠点を補うのが疫学です。

## 新しい疫学研究の可能性

**寺本** フラミンガム研究の特異性はどのようなところにあるのでしょうか。

**磯** フラミンガム研究は内科の研究者が中心ですが、最初から統計学者も入っています。国家プロジェクトとして多彩なバックグラウンドを有する研究者が、循環器疾患の自然歴の探究や危険因子の同定を長期計画にもとづいて行ったのが特徴です。

一般に公衆衛生の研究者は公衆衛生の分野で、臨床の先生はその臨床の分野で、個別的に研究を行ってきました。そのため、医学部の教



桑島 巖 氏



大橋靖雄 氏

育では疫学とか保健統計といったものは学生にとっても非常にとっつきにくかったわけです。ただ、日本でも疫学研究に対する社会的な要請も大きくなってきましたし、状況がだんだん変わってきているように思います。

**桑島** 以前の疫学というのは、断面調査という意味合いが強かったのではないのでしょうか。長期間にわたって追跡するというのはフラミンガム研究以降ではないのでしょうか。

**磯** そうですね。

**桑島** 疫学研究の問題点として、住民の移動が少ないことを前提にしている点ですね。現代は住民が非常に移動しやすくなっていますが、その補完はどのようにやっているのでしょうか。

**磯** 日本で長期間のコホート研究を行っているところは、住民の移動の少ないところが多いですね。大都市の真ん中で行うことは難しい。

**桑島** 企業の健診データは意外と生かされていませんね。定年退職後のフォローが不十分でイベント発生を追うことができないからでしょうか。

**大橋** 一部の企業で行われています。健保のデータを疫学研究として本格的に解析できるとポテンシャルはあると思います。

**寺本** 堀先生がいわれたように、登録がきちんとされていて、フォローアップできて、最終的なエンドポイントまでが追いかけるようなシステムができると、RCTがなくともいろいろな問題点を見いだすことができるはずなのですね。

**大橋** 観察研究の解析方法が、この10年間で急速に進んでいますからできる可能性はあります。疫学研究からの統計手法の開発の一例は、

ロジスティック回帰分析です。これはフラミンガム研究において、Jerome Cornfieldという統計家が発明したものです。最近では観察研究から因果的な効果を推定しようという試みが盛んに行われています。データベースを構築するときには、データの蓄積とともに、それを解析する方法論も開発しなければなりません。

### 疫学と臨床の相互作用から生まれてくるもの

**寺本** 臨床から疫学に対する希望、期待は非常に大きいのですが、疫学の先生たちから臨床医に対して、何かメッセージはありますか。

**磯** 以前ですと臨床研究を行ってしまった後にデータをどう解析したらよいかを、統計家に相談に依っていたのですが、最近は最初から臨床統計家が入るようになり、疫学、統計学の研究者と臨床の先生方が一緒になって互いに乗り入れながら研究を行うようになってきました。そうすることによって、臨床研究から疫学研究へ、疫学研究から臨床研究へというサイクルができあがるのだと思うのです。

**堀** 臨床医と疫学の先生との間にもギャップがあるとすると、私たち臨床医は、疫学からデータがでてきても、メカニズムが説明できないと受け入れないのです。理屈がつけば、臨床医は納得して受け入れます。ですから疫学と臨床の間を何かで埋めないといけない。これはわれわれの仕事ではないかと思えます。

**大橋** 基礎研究者の課題でもありますね。

**寺本** 疫学でわかっていたいくつかの因子に関しても、それを基礎科学の面から解明するとい

うステップが必要なのでしょうね。フラミンガム研究で血清コレステロールと冠動脈疾患の関係が発表されましたが、そのころ臨床からは、リポ蛋白の中の LDL 分画がどうも悪そうだというような実験的なデータがではじめていました。これをうけてフラミンガム研究では LDL が高い人の予後を調べました。また、フラミンガム研究から HDL が防御的に働くことが示されると、今度は HDL の基礎研究がはじまりました。このように互いに高め合っている。疫学がどのように臨床に生きてきているのかを一般臨床の先生方にしめしていくのが、このサイト

の持っている意義なのではないかと思っているのです。大橋先生、最後にコメントをいただけますか。

**大橋** メカニズムを解明しながら、いまだわかっていないところを次の研究で明確にしていく。そういう立体的な構図が見られるとよいですね。

**寺本** 循環器疫学の成果を収載するこのサイトでも、そういう形での動きのある疫学を、ぜひアピールしていただければよいのではないかと思います。本日はどうもありがとうございます。

## interview

# 疫学研究と臨床の交流の場を求めて

epi-c.jp 編集委員

上島 弘嗣 (滋賀医科大学社会医学講座福祉保健医学)

### ●最近の循環器疫学研究の特徴

#### 最近の循環器疫学の動向でこれまでと比べて何か変化がありますか。

一つは国際共同研究も含めて大規模な共同研究が増えてきたことです。これは観察研究だけではなく臨床試験でもそうです。以前は、募集目標数を達成できないなどプロトコルが現実性を伴っていなかった、さらに対象者募集などに対する医師の支援やデータの管理、分析など試験を支える基礎的な体制が整っていなかったために、実を結ばないことも多くありました。しかし、ここ 10 年ぐらいで体制が急速に充実してきたために大規模な共同研究が可能となりました。

もう一つは個人データを統合したメタアナリシスが行われるようになったことです。アジア・オセアニア地域の心血管病の危険因子の解明を目指す Asia Pacific Cohort Studies Collaboration (APCSC) はその代表です。コホートの規模が小さいと多くの詳細な検討、有意差を検出する

のが難しいのですが、個人データを統合することにより有益な所見を見出すことが可能になります。日本でも全国で実施されている既存の循環器コホート研究の個人データを統合する日本動脈硬化縦断研究 (JALS) が数年前から開始されています。今後、こういった大規模な共同研究が増えていくと思います。

#### 食塩と血圧に関する INTERSALT、栄養と血圧に関する INTERMAP など日本が参加している国際共同研究も少なくありません。

INTERSalt は 32 개국 52 集团で行われていますが、国際共同研究は多くの国・施設が参加して、これは既存のデータを利用するのではなく新規にデータの収集を始めるものですから、研究計画から実施に至るまで非常に高度な標準化が求められます。なぜ標準化が大切なのか、デザインを決める理由などを世界中の研究者が理解し、納得してから実施しますので、この過程が非常に教育的なのです。共同研究に参加することが実践的な学習機会になるのです。共同

研究の良いところは参加者の水準が上がっていくことです。次の機会にはその水準以上のものになっていく。そういう意味でも大きく様変わりしてきたと思います。

日本が参加した初めての代表的国際共同研究は1958年に開始されたSeven Countries Study(7か国共同研究)です。その後しばらく途絶えましたが、1982年からのINTERSALT、1996年からのINTERMAP、若年男性の冠動脈石灰化と頸動脈肥厚による動脈硬化と循環器疾患危険因子との関連を日米、韓国で検討するERA-JUMPが2001年、アテローム性動脈硬化の危険因子と治療法の世界規模登録研究であるREACH Registryも数年前に始まりました。

#### ●現在の臨床における過去のデータの意義

**数十年前のデータを解析する疫学研究がリスクの重みに変化している現在の臨床に与えるものは何でしょうか。**

時代が変わっても危険因子の相対危険度そのものはさほど変わらないと思います。例えば、10年前に高血圧が心血管病発症のリスクが2倍だとしたら、現在も2倍です。変わるのは高血圧の罹患数なのです。高血圧患者の絶対人数と割合が減少すれば、高血圧の社会的な影響力は低下します。高血圧の重みが小さくなるということは、人口寄与危険度が低下することではありますが、高血圧が心血管病のリスクであることは変わらないのです。しかし、高血圧の罹患率が低下することにより、それまで高血圧が多かったことにより見えなかった喫煙などのリスクが浮き彫りにされてくるという変化は起こります。そしてこの変わってきた危険因子を臨床試験で検証することが重要だと思います。

昔の状態をみているのではないかという問題があるのは事実ですが、たとえ10年前のデータであっても、そのリスクは発表された時代においても適応できるもので、いつの時代もリスクの検討を観察研究で行うのは大切です、医療政策には必要なデータだと考えています。



上島 弘嗣氏

(2007年1月26日第17回日本疫学学会学術総会会場にて)

#### ●疫学研究と臨床の相互作用

**最近、循環器関連の学会や座談会などに疫学の先生が参加される場合が増えているような気がします。**

実は公衆衛生だけでやっていることに限界を感じ始めたのです。学会などで議論する相手が自分たちしかいないのです。そこで私たちは臨床の学会に出席して発言し始めました。私は高血圧学会や動脈硬化学会などに出かけて行って、「ここがおかしい、あそこもおかしいぞ」などと言い、臨床側に積極的に論議をしかけるようにしたのです。

**それはいつ頃のことですか。臨床の先生方の反応はどうでしたか。**

もうかれこれ20年ぐらい前のことです。高血圧学会でアルコールとイベントの関連の結果を発表した時は、15人ほどの臨床の先生が質問をされるために並ばれたくらい、反応はとても良かったのです。その後も、小規模試験の発表に対してもデザインやデータの不備を指摘し続けました。こういう論議の場ができたことにより試験のレベルが上がっていきました。そのうち生物統計学の先生方も論議するようになり、生物統計学の専門家も疫学研究や臨床試験に参画してもらわなくては行けないと考えようになりました。多分、私は臨床医と最も交流の多い疫学者だと思いますが、これは学会などに積極的に参加したことによるのです。