

# 臨床研究：実戦的基礎知識

国立国際医療研究センター「初期臨床で身につけたい臨床研究のエッセンス」Vol.2 第4章より抜粋・改変

## ポイント

1. 薬の候補物質の、ヒトでの有効性・安全性を検証するために行われるのが臨床試験であり、治験はそのタイプである。
2. 統計学的有意性と臨床的有意性を区別して、クリニカル・エビデンスを評価する必要がある。
3. 近年、臨床研究の規模は拡大する傾向があり、精度確保のためにデータの品質管理・品質保証の重要性が増している。
4. 膨大なデータ量をコンピュータに保存し、複雑さ、多重性に対応するためには、データベースを利用して効率的に管理する必要がある。
5. 生物統計家を含むデータマネジメント部門は、臨床研究の開始前から関与して、臨床医と共同で、データ収集などに関わる必要なルール整備を行う。

一般に‘研究’と名がつくと途端に身構えてしまう人が少なくない。実地診療は、目の前の患者さんの疾病あるいは体の状態の診断、治療、予防を主たる目的としている。臨床医が質の高い診療を行うには、最新の臨床研究の成果を積極的に習得しようとする姿勢が求められる。したがって、臨床医が、自ら臨床研究に従事するというケースよりも、むしろ一流の臨床研究論文を読んで日常診療に役立てようとするケースを先ず念頭におき、本稿では臨床研究の基礎知識を整理してみたい。

## 1. 知る：クリニカル・エビデンス

### 1) エビデンスの検索

Evidence-based medicine (EBM) とは、研究結果からの最善のエビデンス、臨床的な専門技能、および患者の価値観の三者を統合するものである（図1）。後二者には、個々の経験・価値観といった単純に定型化できない要素が含まれるものの、臨床医が妥当な情報を収集し患者に提示することが互いの良好な協力関係を築くうえで不可欠である。この妥当な情報は、基礎医学研究に由来することもあるが、特に患者に焦点を当てた臨床研究に由来することが多い。こうして得られた新しい知見は、従来、受け入れられていた診断検査や治療法を過去のものとして、より正確、有効で安全なものに置き換える。

臨床医が日々直面する病状に対して、最新の情報をどのようにして得るか、その科学的根拠（エビデンス）の度合いはどのくらいと考えられるか、はEBMの実践における重要なポイントである。

患者を診察して临床上の様々な疑問が生じ、それに対する新しい情報を得たいと思う機会は多いであろう。しかし、最新かつ最善のエビデンスに遅れずについていくのは容易でなく、また疑問に対する答えがなかなか見つけられないかも知れない。

それでは、どこで最善のエビデンスを見つけられるのか。临床上の問題点については、教科書に簡便にまとめられているが、頻繁に改訂され適切に管理されて、

系統的かつ批判的に評価された多数のエビデンスを提供してくれる教科書は必ずしも多くない。特定の臨床問題に対するエビデンスは、増加しつつあるオンライン・データベースから見つけられる。

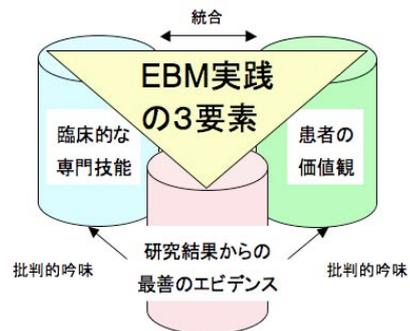


図1. EBM実践の3要素

MEDLINE (<http://medlineplus.gov/>) は、基礎から応用研究に至る生物医学研究全般の、世界最大の文献データベースであり、その作成者である米国医学図書館から検索エンジン PubMed (<http://pubmed.gov/>)によって無料で利用することができる。さらに特定診療分野に絞った、専門的データベースも存在する。もう一つ、包括的な薬学・生物医学文献データベースとして EMBASE が広く知られているが、こちらは Elsevier Science 社による有料サービスである。また科学的価値と臨床的関連性について質の高い論文を選び、提供しているデータベースが ACP Journal Club (<http://www.acpjc.org/>) (旧称 Best Evidence) や Cochrane Database of Systematic Reviews (<http://www.cochrane.org/reviews/>) である。さらに臨床的な興味に関連のある全ての雑誌から、エンドユーザーの代わりに膨大な数のオリジナル文献を収集・解釈・分析し、その質と内容によってエビデンスを選択する明示的な手順に従って作成されたものとして、Evidence-Based Medicine Reviews (EBMR) というデータベースも提供されている (Ovid Technologies 社、有料) (表1)。

データベース名	ウェブアドレス	特徴
MEDLINE	<a href="http://medlineplus.gov/">http://medlineplus.gov/</a>	基礎から応用研究に至る生物医学研究全般の、世界最大の文献データベース。検索エンジンPubMed ( <a href="http://pubmed.gov">http://pubmed.gov</a> )によって無料で利用できる。
The Database of Abstracts of Reviews of Effectiveness (DARE)	<a href="http://www.york.ac.uk/inst/crd/">http://www.york.ac.uk/inst/crd/</a>	医療の介入と方針についての効力に関する系統的なレビューのなかから高品質なものを選出しまとめたもの。
ACP Journal Club	<a href="http://www.acpic.org/">http://www.acpic.org/</a>	科学的価値と臨床的関連性について質の高い論文を選び、提供しているデータベース。
Cochrane Database of Systematic Reviews	<a href="http://www.cochrane.org/reviews/">http://www.cochrane.org/reviews/</a>	治療、予防に関する臨床試験を系統的な方法で吟味し、臨床家に対し現時点での標準的な治療、予防の情報を提供することを目的とした医療の評価調査プロジェクト(コクラン共同計画)がその成果として提供しているもの。最も正確で権威ある治療方針として汎用されている。
EMBASE	<a href="http://www.info.embase.com/">http://www.info.embase.com/</a>	Elsevier Science社による有料サービス。信頼性の高いEBMの情報源4つ (Cochrane Database of Systematic Reviews, DARE, ACP Journal Club, The Cochrane Central Register of Controlled Trials) を1つにまとめたデータベース。
Evidence-Based Medicine Reviews (EBMR)	<a href="http://www.ovid.com/site/index.jsp">http://www.ovid.com/site/index.jsp</a>	Ovid Technologies社による有料サービス。エンドユーザーの代わりに膨大な数のオリジナル文献を収集・解釈・分析し、作成されたもの。

評価済みのエビデンスを収録したデータベース (EBMR など) の活用などで、検索のための情報源は大幅に改良されてきた。臨床判断の裏付けとなるエビデンスを検索する基本的ステップを図2に示す。

## 2) エビデンス・レベル

EBM ではエビデンスを作り出す研究の手法(デザイン)に優劣があるとされている。'エビデンスのピラミッド'(図3)といわれる、研究デザインに基づいた、エビデンスの階層化(エビデンス・レベルの指定)が唱えられている。エビデンス・レベルに関する各研究デザインの概要については他書を参照されたい。

## 2. 知る：臨床研究の種類

### 1) 臨床研究の分類

近年の自然科学、生物学における目覚ましい技術革新とともに、臨床研究は、'知の創造'という学術的役割だけでなく、非常に大きなビジネスともなり、製薬企業やバイオテクノロジー関連企業が大学、その他の学術機関、医療機関と緊密に連携して臨床研究の強力な推進力を発揮することとなった。こうした医薬分野全体の流れのなかで、新規の治療薬、診断手法の臨床応用を目指した臨床試験が大きな比重を占めるようになってきた。

一般に、日々の臨床的観察および基礎的研究の成果を、疾病の新しい治療・診断手法として臨床へ導入するために臨床研究は行われる。ひと口に臨床研究といっても多様な種類の研究がなされており、図4に示す通り、(1) 臨床試験(次項参照)からアウトカム研究、医療の質研究、医療決断分析に至る、主に治療(診療)内容の検証を目指したもの、(2) 臨床知識の発見を目

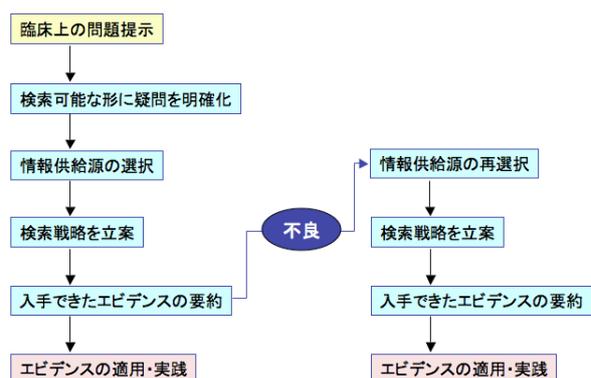


図2. エビデンス検索の基本ステップ

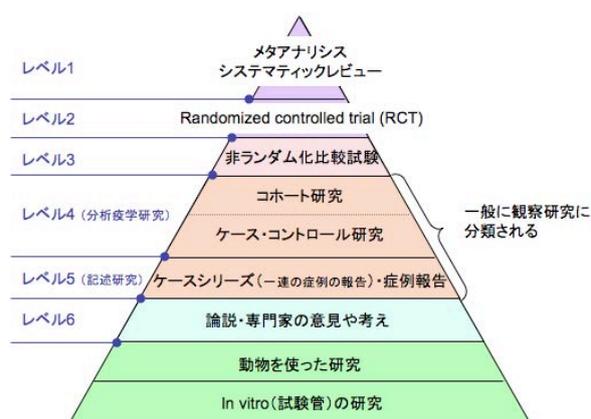


図3. エビデンスのピラミッド

指した臨床疫学研究や社会心理学研究、さらに(3) トランスレーショナル研究を通じた基礎的研究との連携に関するもの、に大きく分類される(広義の臨床研究、[コラム参照](#))。

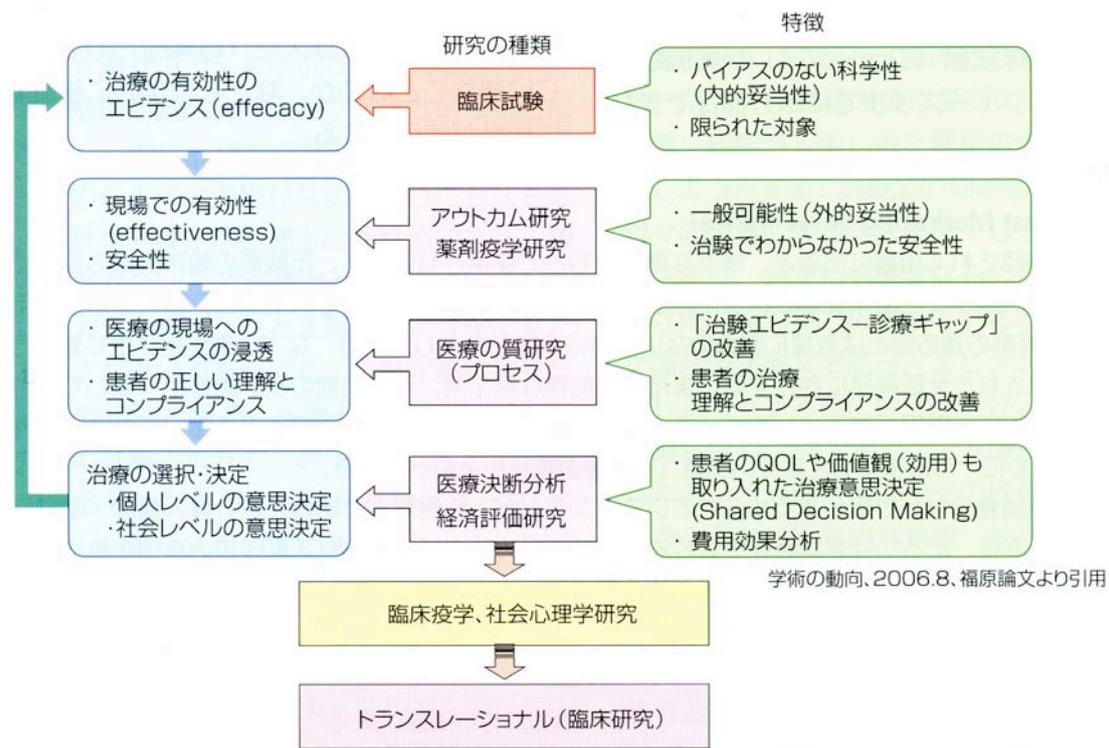


図 4. 臨床試験の種類

コラム：広義の臨床研究

臨床研究は、医学の最新の研究知見を検証する臨床試験と同義と考えられがちである (狭義の臨床研究)。しかし広義の臨床研究が包含する研究領域はもっと広く、臨床試験に加えて、診療レベル、社会レベルにおける様々なステップの研究が必要とされる。薬剤疫学研究 (医薬品の安全性の検証)、アウトカム研究 (実際の診療現場におけるエビデンスの確立)、医療の質評価研究 (エビデンスと実地診療のギャップの改善、及び患者の理解とコンプライアンスの改善)、医療の経済性評価・医療政策に関する研究 (費用効果分析)、医療決断分析研究 (治療の選択・意思決定に関する分析)、医療情報やコミュニケーションに関する研究、なども広義の臨床研究に相当する。

2) 臨床試験と治験

臨床研究のなかで、薬の候補物質の、ヒトにおける有効性・安全性を科学的に検証するために行われるのが臨床試験 (Clinical Trial/Study) である。そのなかで、製薬企業等が医薬品としての販売許可を厚生労働省から取得するためのデータ収集を目的として実施する臨床試験のことを「治験」と呼ぶ。2002年に薬事法が改正され、それに伴う改正GCP (医薬品の臨床試験の実施の基準) の施行によって2003年7月30日より医師・医療機関の主導による治験の実施が可能となった。すなわち、臨床医は、製薬企業が実施する治験だけでなく、医師主導型の治験とそのデータマネージメントにも携わる機会がでてきたわけである。



図 5. 治験のフェーズ

薬の候補物質は、最初に動物で効果や毒性についての試験 (「非臨床試験」) を実施し、その効果と毒性の程度を確認してから、「治験」に進む。治験には以下の3段階 (フェーズと呼ばれる) があり、各段階での安全性や有効性を順に確認しながら開発が進められる。

第1相 (フェーズ1)

少人数の健康成人男子で、ごく少量から薬の候補物質の投与量を少しずつ増やしていき、安全性の確認と、物質の体内動態 (Absorption 吸収、Distribution 分布、Metabolism 代謝、Excretion 排泄) を評価する。

第2相 (フェーズ2)

薬の候補物質が効果を示すと予想される比較的少人数の患者で、有効性・安全性と投与量・投与間隔などの関係を探索する。

### 第3相（フェーズ3）

多数の患者を対象に、第2相で推定された用量・用法によって既存の類似薬ないしプラセボ（偽薬）との比較を行い、薬の候補物質に関する有効性と安全性、使い方を確認する。比較に際しては、客観性を担保するために無作為化（Randomization）され、盲験下で治験が行われることが一般的である。

「治験」として実施した各フェーズの結果は、新薬承認申請の際、製薬企業が厚生労働省に提出する資料になる。そして審査センター（独立行政法人 医薬品医療機器総合機構）での審議を経て承認された場合に、医薬品として販売することが可能となる。この承認に際して、確実な有効性と安全性の保証がなくても医薬品として認可されるケースもある。特に、現在、有効な治療法がなく致死的な疾病に対しては、リスクとベネフィットのバランスを考慮したうえで承認の是非が判断され得る。

製薬企業が医薬品の市販を開始した後は市販後調査（PMS：Post Marketing Surveillance、[コラム](#)参照）が義務付けられ、そのデータを基に再審査、再評価が行われる仕組みとなっている。市販後調査の一つとして市販後臨床試験があり、治療的使用に関しては第4相臨床試験（フェーズ4）が実施される。この段階で、フェーズ3までに十分捕捉できなかった副作用情報や、経験のない併用薬による安全性の評価、などが行われる。綿密な市販前臨床試験（治験）を経て有効性と安全性が確認されたはずの薬剤が、市販後に重篤な有害事象のため販売中止になることがある。これは、臨床研究のゴールドスタンダードとされるRCTが万能ではないからである。すなわち、RCTは薬剤の効果判定には非常に有用であるが、有害事象の検出には不十分であり、しばしば不適切である。

#### コラム：市販後調査（Post Marketing Surveillance）

医薬品は、承認されて市場に出ると、様々な環境のもとで使用され始める。市販前の臨床試験の段階では十分に検討されていなかった臨床現場の状況、たとえば、高齢者や子供に投与された場合、他の薬剤と併用された場合、（当該薬剤の適応症とは直接に関係のない）他の疾患を合併した場合、などに副作用が生じ易いという可能性は、承認された全医薬品に存在する。また比較的稀な副作用、潜伏期間の長い副作用に関しては、市販前臨床試験での検出に限界がある。したがって、臨床試験で効果の証明された医薬品であっても、安全性をさらに確立するために、市販後にも注意深い調査（市販後調査）が必要とされる。

この市販後調査のなかで最もよく知られているシステムが、自発報告である。これは、医師、薬剤師といった当事者から直接、あるいは製薬会社を通じて（間接的に）政府に報告される副作用の事例を集めたものである。しかし、自発報告だけで全ての事例をカバーできないため、観察研究（特に薬剤疫学研究）での系統的モニタリングも必要とされる。

### 3. 臨床研究における結果の解析・解釈

生物の行動、反応などにはすべて“ばらつき”がある。この“ばらつき”を認めながら、目の前で観察された反応（たとえば介入に伴うイベント発生の低下）が似たような人々（母集団）のなかでどれくらい稀なことなのか、などを科学的に検証するために統計が使用される。

#### 1) 統計学的有意性と臨床的有意性

臨床現場では、目の前にいる自分の患者に対して「ある介入（たとえば投薬治療）を行うべきか、従来通りの介入にとどめるべきか」の二者択一の判断を迫られる。こうした臨床上の疑問点に答えるべく RCT などの臨床研究が実施され、介入群と非介入群との間で、エンドポイントとして設定されたイベント発生率の違いが「偶然のばらつき」によって説明できる範囲に収まっているかどうか（収まっているという仮説を帰無仮説という）、が統計学的に検定される。この「偶然のばらつき」の範囲を、母集団から研究対象群をランダムに100回抽出した際に95回は収まる範囲と決め、それを超えた場合を、危険率ないし有意水準5%（ $p=0.05$ ）で帰無仮説が棄却されたという。

しかし、 $p$  値が0.05より小さいからといって、すぐにその結果が臨床的に重要であるということにはならないし、逆に $p$  値が0.05より大きいからといって、その介入には効果がないということにもならない。統計学的検定は数理上の計算値であり、標本数などの条件によって結果が大きく影響される。標本数さえ大きければ、たとえ臨床的には役に立たないほど小さな違いであっても、統計学的には有意差があると判定され得る。 $p$  値の大きさではなく、むしろ介入の効果の大きさ（effect size）が臨床的に意味のあるものかどうかを検討せねばならない。この目的で、相対リスクの推定値である、Cox 比例ハザードモデルの相対ハザードや多重ロジスティック回帰分析におけるオッズ比が汎用される（[コラム](#)参照）。

一方、標本数が少ないため、本来は臨床的に意義があるのに統計学的に有意差が認められていないケースは実際、極めて多い。このような事態を回避すべく、臨床研究の計画段階で、必要な標本数を推定するためのパワー計算を行う必要がある。詳細は他稿に譲るが、パワー計算における統計学的検出力を規定する要素として、 $\alpha$ エラー（本当は差がないのに、差があるとしてしまう誤り：偽陽性に相当）と $\beta$ エラー（本当は差があるのに、差がないとしてしまう誤り：偽陰性に相当）、そして臨床的に意義があると考えられる effect size の推定値、の3つが重要である。たとえば、対照群で40%の患者にイベント発生が予測される疾患に対して、新しい治療法の効果を調べる際、その適用にてイベント発生率が30%以下になれば、臨床的に意義があると考えられる場合（相対的減少率25%）、 $\alpha$ エラー5%、 $\beta$ エラー20%の設定では各群380人の被験者が必要と計算される。

コラム：オッズ比

オッズ比 (Odds ratio : OR) は、ある事象の起こりやすさを2つの群間で比較して示す統計学的な尺度である。このオッズとは、確率論だけでなくギャンブルでも盛んに使われてきた数値で、ある事象の起こる確率を p とした場合、 $p/(1-p)$  のことである。当該事象の、2群における確率を p (第1群)、q (第2群) とすれば、オッズ比は、

$$\frac{p / (1-p)}{q / (1-q)} = \frac{p \cdot (1-q)}{q \cdot (1-p)}$$

ここでオッズ比が 1 とは、対象とする条件あるいは事象の起こりやすさが両群で同じということであり、1より大きい (小さい) とは、条件あるいは事象が第1群 (第2群) でより起こりやすいということである。

2) エビデンスの適用

臨床医は、主観的・直感的ではあるが、身体 (理学) 所見や検査結果等に基づいて鑑別診断を行い、知識・経験や入手し得た最新の情報に基づいて (適用しようとする) 治療法の効き易さを予測しながら、個々の患者に役立つ決断を下している。しかし、この決断は、必ずしも確実なものではなく、大まかに言えば確率論的の見方に従っているともいえる。臨床医が臨床医学論文を読み、エビデンスを実地診療に適用しようとする際には「患者が□□という病気である確率は○○%であり、この病気の患者に対して、△△という治療法は○%の確率で効果がある」というような「確率的な情報」を得ようとしているのである。統計は、この「確率的な情報」さらには effect size を知るために必要とされる手段である。論文中に報告された研究対象の患者群と、決断を下そうとしている目の前の患者との間には様々な面での違いがあるとしても、補正を加えることで自分の患者にエビデンスの適用が可能だと判断した場合、ベイズの定理 (コラム参照) を応用して

エビデンスを各々の診断と治療の意思決定に役立てていくことになる。

『確率論を巧みに用いることで不確実性の数量化を行うことが可能なことを示した』のがベイズの定理である。医療においては、事前確率と事後確率の考え方をを用いることで、診断 (検査結果に基づく検査後確率の推定) や RCT での中間解析・モニタリングに応用されている。

コラム：ベイズの定理

1763年にイギリスの牧師、トーマス・ベイズによって発見された確率論の定理であり、いくつかの未観測要素を含む、コンピュータでの推論等に応用される。何が (今後) 起こる可能性はその事柄の過去の発生頻度を使ってほぼ推測できるという考え方のもと、その確率計算に用いられる。

すなわち、

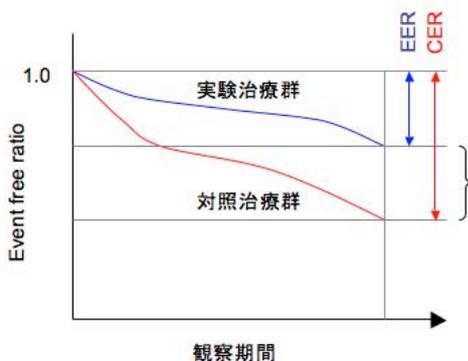
$$P(B) = \text{事象Bが発生する確率 (事前確率)}$$

$$P(B|A) = \text{事象Aが起きた後での、事象Bの確率 (事後確率ないし条件付確率)}$$

とした場合の、 $P(B|A)$ の算出に用いられる。

たとえば、ある年齢で癌である確率  $P(B_1)$  と癌でない確率  $P(B_2)$ 、癌である人が検査で陽性となる確率  $P(A|B_1)$  と、癌でない人が検査で陽性となる確率  $P(A|B_2)$  を、調査結果をもとにして求めておけば、この検査を行って陽性と出た人が癌である確率は  $P(B_1|A)$  によって求められるのである。そもそも、癌であるか否かということが、陽性かどうかという検査結果の確率を決めているものの、ベイズの定理は、更に検査結果からみて癌であるかどうかの確率を求めている (推測している) ことになる。

治療に関するエビデンスの適用としては、RCTの結果の解釈が重要である。その effect size は、用いる効果指標により印象が大きく異なる。RCTでは治療群でアウトカム (転帰) の発生の低下が生ずるかどうかが



EER: 実験治療時イベント発生率  
CER: 対照治療時イベント発生率

$$\text{CER} - \text{EER} = \text{ARR (absolute risk reduction)}$$

= 絶対リスク減少率  
= 実験治療の「お蔭で」イベントを起こさずに済んだ人の割合

$$\frac{1}{\text{ARR}} = \text{NNT (number needed to treat)}$$

= 治療必要数  
= 何人の患者が同じ治療を受けると、1人のイベントが予防できるか

$$\text{相対リスク減少率 (relative risk reduction: RRR)} = (\text{CER} - \text{EER}) / \text{CER}$$

$$[= \text{ARR} / \text{CER}]$$

【留意点】

結果の大きさは、生存曲線で表すことが最も適している。  
他の指標を用いる場合、RRRで表されると、一般に、より高い効果があるかのような印象をもつ可能性が高い。

図 6. 臨床研究 (RCT) における効果指標

表 2. 臨床研究のタイプと統計解析のパターン

タイプ	概要	種類	解析方法(1)	解析方法(2)	解析方法(3)
検証的研究: 多くのアウトカム・スタディ	臨床的イベント発生の有無をアウトカムとして時間的経過を比較したもの	RCT、コホート研究	生存率推定: Kaplan-Meier法	治療群間の生存率の差の検定: Log-rank検定	対照群に対する、実験群の相対リスクの推定: Coxの比例ハザードモデル
検証的研究: イベント発生頻度差の検証など	イベント発生の有無をエンドポイントとしながらも時間的経過を考慮しないもの	RCT	イベント発生頻度の差の検定:Fisherの直接確率検定、相対リスクの推定:Mantel- $\chi^2$ 乗検定、Mantel-Haenzelの相対リスク検定	イベント発生率の用量-反応関係の検定: 拡張Mantel検定、Cochran-Armitage検定	参照群に対する、各群の相対リスクの推定: 多重ロジスティック回帰分析、Coxの比例ハザードモデル
検証的研究: 特定曝露因子の関与など		コホート研究 (段階的なトレンドの検定)	イベント発生率の推定: (複合)エンドポイントの認定		
探索的研究	注目するイベントの危険因子・予測因子を探索するもの	コホート研究	要因の分析:多重ロジスティック回帰分析、Coxの比例ハザードモデル、重回帰分析		
検証的/探索的研究	低発生率のイベントと曝露要因との関係、または前向き研究の必要性の検討	症例対照研究	条件付きロジスティック回帰分析		

を検討し、その効果指標として、RRR（相対リスク減少率）、ARR（絶対リスク減少率）とNNT（治療必要数）の3つが用いられる（図6）。RRRは概念的に受け入れ易いが、effect sizeをかなり大きく、ときに誇張して見せる印象がある。むしろ、‘実験治療の真の恩恵を受ける人の割合’を示すARR、あるいはその逆数であり‘何人の患者に実験治療を行えばその治療効果により1人の運命を変えられるのか’を示すNNTで表現した方がよいとされる。

### 3) 臨床的仮説と研究のタイプ

研究者がある疾病の有病割合や発生率を調べようとする場合、それらを知りたい集団のごく一部を観察することで全体を推測せざるを得ないことが多い。注目する集団から被験者を抽出する手順はサンプリングと呼ばれ、その方法により、疫学研究の方法は3つ—サーベイ、コホート研究、そしてコホート研究の特殊形ともいえるケース・コントロール研究（症例対照研究）—に分類される。さらにこれらとは別枠になるが、RCTなどの臨床試験は、コホート研究といと同土のような関係にある。

サーベイでは、有病割合などの、ある一時点で調べられた個々人の状態を記述する。母集団全体に対して行うには手間と費用がかかりすぎるような測定を対象集団で行うものであり、横断研究とほぼ同義である。また、疾病発症の危険因子、予測因子などを探索するためには疫学研究のなかに時間の概念を導入する必要があり、その最も直接的な形態がコホート研究である。通常は、地域住民対象のコホート研究が行われるが、病院を受診した患者を対象とした病院コホート研究も可能である。比較的ありふれた疾病等に関する研究には、コホート研究はよく適するが、短期間では発生しないような稀な疾病等を観察するには、コホート研究は適していない。多数の対象者（数千人とか数万

人レベル）を長い年月にわたって追跡せねばならないからである。そこで代わりに症例対照研究が実施される。その一つとしてコホート内症例対照研究（nested case-control study）と呼ばれる方法も見られる。これは、もともと別の目的でコホートとして追跡されている人々のなかで、あるイベントが発生し、それに何らかの曝露要因が関与するかどうかを、遡って調査する場合などに用いられる。

臨床研究は、特定の臨床的仮説をたてて、それを検証するために実施されるのであり、仮説の種類・内容によって研究のタイプと統計解析の手法は自ずと規定される。臨床医にとってインパクトの大きな臨床研究では、①臨床的イベント発生の有無をアウトカムとして時間的経過を比較したもの、②同じくイベント発生の有無をエンドポイントとしながらも時間的経過を考慮しないもの、そして③注目するイベントの危険因子・予測因子を探索するもの、の3タイプが主要部分を占める。各々のタイプの臨床研究では、表2に示すような一定のパターンで統計解析が行われる傾向がある。

### 4. 臨床研究での留意点

臨床研究を進める際に重要な過程として、プロトコール作成、データ収集・管理、そして倫理的配慮が挙げられる。これらは、臨床医が自ら臨床研究や治療に関与する場合はもちろんのこと、臨床医学の論文を読み、その研究結果を（自分たちの）目の患者にどのくらい適用できるかを判断する場合にも役に立つものである。以下、プロトコール作成およびデータベースについて、簡単に留意点を述べる。

#### 1) プロトコール作成

プロトコール（protocol）とは「複数の者が対象となる事項を確実に実行するための手順等について定

めたもの」で、規定、議定書などと訳される。特に情報工学では、コンピュータ等の電子機器間で通信する際の取り決めを指す。臨床研究の開始にあたって、研究者は先ず“臨床研究プロトコル（以下、プロトコル）”を作成する。

プロトコルには、誰が主任研究者（principal investigator：PI）であるか（責任の所在）、研究で何が検証されるべきか（臨床的仮説の設定）、どのようになされるべきか（研究デザイン）、計画の進捗状況がどのように評価されるか（有害事象・有効性モニタリング）、最終的なデータはどのように解析されるか（臨床・生物統計解析）、などが記載される。なかでも研究デザインは、研究を完遂させる（臨床的仮説を適切に検証する）ための科学的側面だけでなく、患者ケアの倫理的側面も十分に考慮せねばならない。プロトコル作成上の留意点を表3に記す。詳細については他稿を参照されたい。

表3. 臨床研究プロトコル作成上の留意点

- 1 書類作成の実務的事項：担当者・事務局の設置
- 2 研究デザイン：研究目的への適合性
- 3 研究デザイン：科学的側面
- 4 研究デザイン：患者ケア
- 5 研究に関する責任：法律上の責任
- 6 研究に関する責任：有害事象の報告に関する責任
- 7 研究に関する責任：実施における様々な局面での責任

主任研究者は、先ず、対象となる研究分野を選択して、当該分野でどのような成果が既に発表されているのかを網羅的に検討し、さらに何が有望なテーマであるかを見つけ出さねばならない。そのうえで、注目するテーマについて綿密に研究計画を立て（ここにプロトコル作成の作業が相当する）、どれだけの研究費が必要かを検討し、その資金を得るためにgrant申請をすることとなる。一般にプロトコル作成の初期段階は、主任研究者または研究事務局（Study coordinator）によってなされるが、その後、共同研究者とも分担しあいながら入念にプロトコルの作成を進めてIRB（institutional review board、機関審査委員会）等での審査・承認を受けていく。

## 2) データベース

臨床研究で収集されるデータには様々な種類がある。被験者の性別、年齢、病名などの背景データ、薬物履歴、手術歴などの治療法に関するデータ、そして臨床検査データなど多岐にわたるため、臨床研究の目的に応じてデータの分類法を検討せ

	ファイルシステム	データベースシステム
データへのアクセス	固定的	多角的
応用プログラムとの関係	ファイル構造の変更には、応用プログラムの変更が必要	応用プログラムとは独立にデータベース構造を変更可能
データの重複・一貫性	同一データが複数ファイルに重複して含まれ、データ間の矛盾が生じ易い	データの重複がなく、一貫性が保たれる
データの維持・管理	複数ファイル全体の総合的管理は難しい	データ全体が一括管理される

表4. ファイルシステムとデータベースシステムの違い

ねばならない。

臨床研究のデータは、一症例分であっても複雑な‘構造’をもっている。たとえば、同一の検体検査情報を繰り返し収集するケースもあれば、薬物履歴のように追跡調査中に種類・量が変動するケースもある。これらの膨大なデータ量をコンピュータに保存し、複雑さ、多重性（同一データが複数の場所に保管されること）に対応するためには、何らかのデータベースを利用して効率的に管理することが妥当である。データベースの構成要素は図7に示す4つである。

### データベースの4つの構成要素

- ① ハードウェア：記憶媒体上にあるデータベース機器本体
- ② 基本ソフトウェア：データベースを体系的に管理するためのプログラム
- ③ 応用ソフトウェア：データベースを応用するためのプログラム
- ④ その他：上記の3要素を扱う人間

図7. データベースの4つの構成要素

上記の応用ソフトウェアとは、利用者サイドから見たデータベースの利用方法に関する部分である。利用者は、業務用のプログラムを作成する応用プログラマーとエンドユーザーとに分けられる。通常の、データベースの利用者はエンドユーザーである。

データベースシステムに対峙するのがファイルシステムであり、両者の違いのポイントを表4にまとめる。

## 5. 臨床研究におけるデータ収集・管理

ある病気の患者さんに対して、例えば異なる2種類の治療法のいずれが「より有効か（安全か）」を検証したい場合、2群間での治療反応性を比較する必要がある。日常診療では、個別医師が任意のタイミングで、個々に必要と判断した検査を行っており、検査データのなかには‘ノイズ’（絶食条件が守られていない空腹時血糖など）が混入している可能性も高い。しかし、研究という観点からすれば、一定のタイミングで、一定の種類の臨床データを系統的に収集していかねばならない。

こうした意図的なデータ収集を行ううえで、品質管理・品質保証、収集の手続きとデータベース作成等の知識が不可欠であり、その概略について説明を加える。

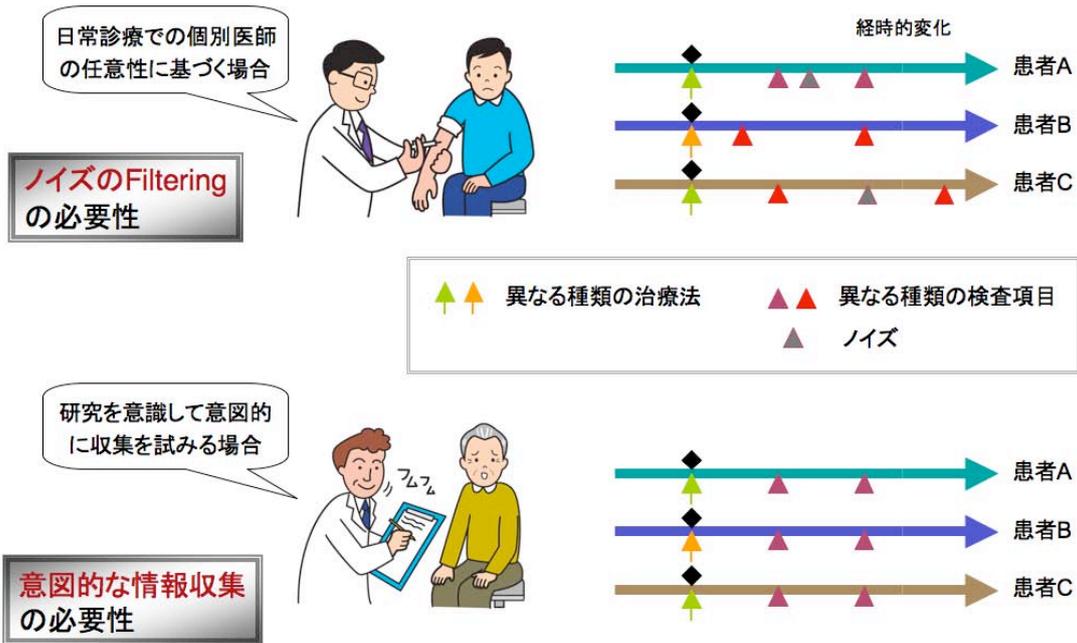


図8. 日常診療で得られるデータと研究用データの違い

### 1) 臨床研究の大規模化

臨床研究（治験を含む）において、その規模を決定するのは研究費の額であるが、それを別とすれば、検証しようとする仮説とその研究デザインである。比較的少ない被験者数を対象として、限られた量の臨床情報のみを用いて実施できるものから、数千～数万人を対象として、広範な臨床情報に関する追跡調査が必要となるものまで様々である。

近年、研究成果としてのエビデンスに求められる「客観性」の水準が高くなるにつれて、臨床研究の規模は拡大する傾向にある。その典型的な例として、高脂血症治療薬（スタチン）の、主要な「偽薬-対照臨床試験」の総被験者数（と糖尿病患者数）を図9に示す。

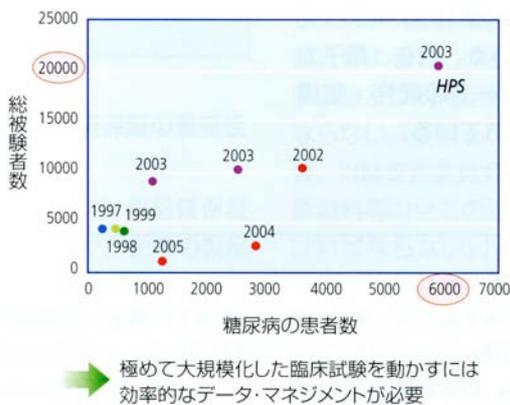


図9. スタチン(高脂血症治療薬)の主要な「偽薬-対照臨床試験」

大規模臨床試験が目目され始めた約10年前には総

被験者数4000-5000名(糖尿病患者数としては500名以下)のレベルであったが、2003年のMRC/BHF Heart Protection Study of cholesterol-lowering with simvastatin (HPS)では総被験者数20000名(糖尿病患者数としては約6000名)のレベルにまで大規模化した。これだけの規模となると、単独の施設、少人数の研究者で稼働させていくことは実質的に困難であり、多施設共同研究として、効率的なデータマネジメント機能を備えた研究組織を構築する必要がある。

### 2) 臨床研究におけるデータマネジメント業務

臨床データマネジメントという業務は多岐にわたり、一部分を除いて臨床医が直接に従事することは少ない。しかし、臨床試験（治験）に責任医師として参画する場合、および臨床研究を自ら企画・立案、遂行しようとする場合には、連携して行われるデータマネジメント業務に関する基礎知識を身につけておく必要がある。

臨床研究・治験のデータを管理すること、なかでも品質管理・品質保証を実施することがデータマネジメントの本質的な業務である。従来、紙ベースの症例報告書を用いて臨床研究・治験が行われていた時代には、その終了後にデータマネジメントの業務のほとんどが集中していた。しかし、データの品質管理・品質保証のレベルを向上させるためには、臨床研究・治験の開始前からデータマネジメント部門が関与して、臨床医と共同で、必要なルール整備がなされるべきこと、業務を分散させて集中的なオーバーワークを防止すること、などが求められている。

## 6. データの階層性

臨床研究のデータは、一症例分であっても複雑な「構造」をもっている。たとえば、同一の検体検査情報を繰り返し収集するケースもあれば、薬物履歴のように追跡調査中に種類・量の変動するケースもある。これらの膨大なデータ量をコンピュータに保存し、複雑さ、多重性（同一データが複数の場所に保管されること）に対応するためには、何らかのデータベースを利用して効率的に管理することが妥当である。

データベースは、対象となる事象（この場合、臨床研究に関わる医療情報）を、モデル（データモデルという）を使って抽象化し、コンピュータ上に投影したものと見ることができる。現在、コンピュータで利用可能なデータベースのモデルとしては、図10に挙げた3つが代表的なものである。詳細に関しては他書に譲るが、各々について簡単に説明を加える。

1) 階層型データベース：段階的に細分化できるデータの管理に適しており「木」構造を示す。親レコードは複数の子レコードを持つことができるが、子レコードからは親レコードは1つに限られるという特徴をもつ。データベース作成後のデータ構造の変更が難しく、それを解決するために1970年代からリレーショナル・データベースが開発されてきた。

2) ネットワーク型データベース：階層型データベースには、子レコードから複数の親レコードと関連付けられないという制限があるのに対し、複数の親レコードと関連付けられるように改良したものである。レコードとレコードは網の目のように結ばれる。

3) リレーショナル・データベース：現在「データベース」と言えば、リレーショナル・データベースを指すと言っても過言ではないほど普及している。データは、行（レコード）と列（フィールド）から成る2次元の表（テーブル）として構成される。複雑な構造のデータを単純化して取扱い、イメージしやすく、データ構造の変更や拡張がしやすい等の特徴がある。図10に示した通り、適宜作成されたテーブルどうしをキー項目（この場合は、検体番号）により組み合わせることができる。

臨床研究のデータには構造的な違いだけでなく、研

究ごとに収集される次元（ないし階層）が異なるという面もある。薬効評価を目的とした場合でも、通常、たった一つの臨床試験・治験を実施して終わることは少ない。特定の薬剤に関する臨床試験・治験に関しても複数の（施設ないしサブテーマで）臨床試験が実施されることが一般的であり、これらを合わせてデータの集計や解析を行う必要が生じてくる。したがって、関連するデータ収集基盤は、いくつかの「階層」に分けて考えた方が効率的に処理することができる。たとえば、全ての臨床試験・治験（薬剤の種類に無関係）に共通して用いられるようなデータを「共通レベル」、特定の薬剤に対して用いられるデータを「薬剤レベル」、（当該薬剤の）個別のプロトコルに対して用いられるデータを「プロトコル・レベル」、さらに（多施設研究において）施設単位で用いられるデータを「施設レベル」、患者ごとに用いられるデータを「患者レベル」と分類することができる。一つの臨床試験・治験のみを考えれば、収集データのコードを自由に定義してもよい（たとえば男性を「1」、女性を「2」と定義する）が、データを併合して集計・解析するためには「階層」を考慮したうえでコードの統一を図る必要が生じる。

我が国で行われる臨床試験・治験において利用可能なコードは、未だ十分に標準化されていない。しかし、いくつかの分野で代表的に使用されているコードはあり、たとえば、疾患名に関しては国際疾病分類（ICD：International Classification of Disease and Related Health Problems）が1995年より死因などの統計分類に用いられている。現在、電子カルテの普及に向けて、標準化コードの開発・整備が精力的に進められている。

### トピックス：ゲノム時代の臨床研究

21世紀はバイオ・ゲノムの世紀だといわれている。過去数年間だけをみても、目覚ましいスピードで解析技術は進歩し、それに伴って膨大な研究成果が蓄積してきた。こうしたゲノム科学の発展を先導したのが、1990年から15年間の計画として始まったヒトゲノム計画（Human Genome Project：HGP）である。ヒトのゲノム（DNAにコードされた全遺伝情報であ

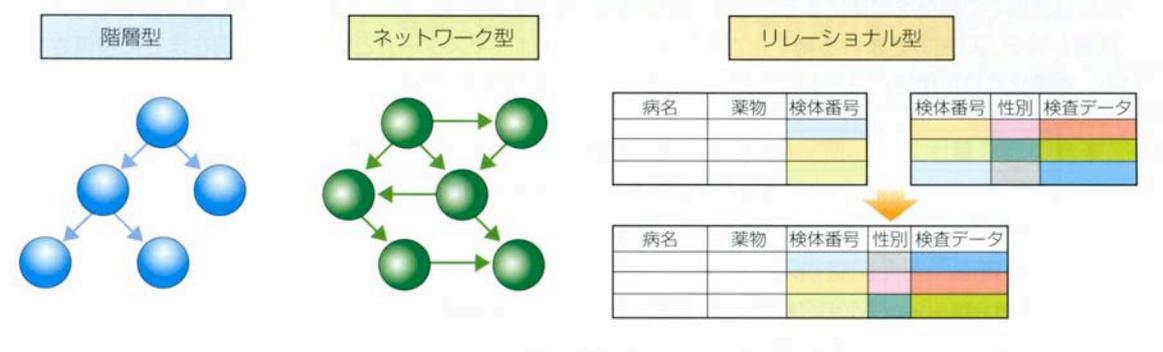


図10. データベースの種類

り、遺伝子コード配列と非コード配列とから成る)の全塩基配列を解読しようという同プロジェクトは、従来の医学・生物学の「常識」を大きく塗り替えた。単純な DNA 塩基配列の中から、遺伝子の境界を決定したり、生物学的な特徴を予測したりする作業(アノテーションと呼ばれる)の必要性が、HGP を契機に大きくなり、バイオインフォマティクスという研究分野が注目を集めることとなった。

こうしたゲノム科学の変化により、病気の克服を目指す医学研究においても大きなパラダイムシフトが生じている。これまでの分子生物学的研究の主流は、特定の遺伝子から出発し、それと関わりのある因子群を調べて、「点と線」を徐々に広げていくアプローチであった。しかし、HGP の推進とともに発展した high-throughput technology は、出発点(ないし着目点)を特に定めることなく、全ゲノムからの“探索的アプローチ”を可能とした。その代表的な解析技術が genome-wide association study (GWAS)における大規模 SNP typing と、トランスクリプトーム解析における DNA マイクロアレイである(コラム参照)。

また、医療の大きな潮流という視点では、集団として標準化された手技・手法を確立し、その精度を高めるとともに、この“万人向け”医療を“個別化”することが大きな課題であろうと考えられるようになってきた。たとえば、高血圧、糖尿病や心筋梗塞などの生活習慣病に関して、「どの治療(薬)が最も有効であるか」という科学的選択を行うには、個々人の遺伝素因、生活習慣に関する情報が不可欠となる。個々人の治療を考えた場合、食事療法(減塩、カロリー制限)や運動療法が有効なのか、特定の種類の薬物が効き易いのか、などは遺伝的に規定された体質に密接に関係する。すなわち、ゲノム科学の進歩は、心血管疾患やそのリスクファクターのような、ありふれた複合疾患の治療法と予防法を更に改良してくれるものと期待され、そのためにゲノム情報をも取り入れた形の臨床研究が(徐々にではあるが)推進されつつある。

HGP に先導された医学研究のパラダイムシフト、“hypothesis-driven approach”から“data-driven approach”への変遷は、疾病の理解と、それに基づく創薬のシーズ提供という面で、大きな突破口をもたらす可能性がある。その一方で、遺伝子診断という面でも、ゲノム科学は相当な医療変革をもたらすと考えられる。個々の疾病に果たしてどれくらいの数の病因遺伝子が存在するのかは未だ不明であるが、一部の例外を除き、病因遺伝子各々の効果は、相対危険度として 1.1-1.5 倍程度の上昇と推定されている。しかし、それらの遺伝情報を組み合わせることによって、高リスク群(一般集団中の平均よりも 3~7 倍リスクが上昇するという試算がある)を同定することができ、その人々に集中した、効率的な治療・予防介入戦略の策定に役立ち、その結果、住民レベルでの保健医療にとっては大きな経済的メリットがもたらされる。

また、がん臨床においては、がん組織の特性を遺伝子レベルで調べ、それに基づいた治療薬(分子標的薬)の選択(ゲノム情報の活用)が既に現実のものとなり

つつある。

ただし、こうした技術が臨床現場に導入される際には、一次医療を施す人々が関連する知識を身につけ、医療を受ける人々の権利確保(プライバシー侵害や遺伝子差別の防止)などに十分な注意が払われねばならない。

#### コラム：全ゲノム関連研究と DNA マイクロアレイ

メンデル遺伝の疾患遺伝子を探索・同定するための方法は、大まかに候補遺伝子アプローチ(順行遺伝学)とゲノムスキャン(逆行遺伝学)に分類されることが多い。前者が何らかの‘機能情報’に依存するのに対して、後者はゲノム上の‘位置情報’に依存するアプローチである。またポジショナル・クローニングというのは、ゲノムスキャン等を通じて見出された、おおよその染色体位置に関する情報のみを頼りにして、疾患遺伝子を同定する方法である。

特定の集団において、ある遺伝的バリエーション(多型)が、注目する疾病表現型(フェノタイプ)と関連するか否かを検証するのが関連研究(association study)である。ここでいうところのフェノタイプには量的(quantitative)なもの(血糖値など)と質的(qualitative)なもの(糖尿病の有無など)とがあり、前者については集団全体を対象とし、後者については(集団の一部である)case 群と control 群の間の比較により、統計学的有意性を調べる。過去においては、殆どの association study は候補遺伝子アプローチとして実施されてきた。一方、ゲノムスキャンでは、全く仮説をたてずに「探索する」ため、従来は知られていなかったパスウェイに関わる遺伝子を同定できる可能性がある。遺伝的インパクトが比較的マイルドな多因子遺伝疾患のゲノムスキャンにおいて、成否の鍵を握るのは、全ゲノムの多型(特に SNPs)情報の網羅性向上と、それらを代表できる‘タグ’SNPs(遺伝子多型のロードマップの役割を果たす DNA 配列の一部分)の選出である。ヒトゲノムは 32 億塩基対から成り、そこに約 1500 万個の SNPs が存在すると推定されている。これら全ての SNPs を一人ひとり調べることは大変な作業であるが、タグ SNPs を中心に調べることで省力化、低コスト化が可能となる。このように絞って 50 万 SNPs を typing することで、計算上は、全ゲノムの約 90% をカバーできる。

ヒトゲノム上の遺伝子総数は 2~2.5 万と推測されている。これらの遺伝子(のトランスクリプト)の発現を一度に調べられる技術が DNA マイクロアレイ(DNA チップ)である。もともと、スポッティング装置により基板上に DNA を整列(アレイ)させ、固定したものを DNA マイクロアレイといい、光リソグラフィ技術を応用して基板上でオリゴ DNA を *in situ* 合成したものを DNA チップと名付けたが、最近では両者を同義で用いることが多い。

文責：加藤 規弘