

# 医学研究データの社会的流通と研究倫理†

大橋 靖 雄\*

## Public Dissemination and Ethical Aspects of Medical Research Data

Yasuo Ohashi\*

医学に関係する情報の社会的流通と医学研究倫理の関係について、三つの事例を取り上げて問題を提起する。薬物や新しい治療法の有効性を調べるには、人間を対象とする臨床試験が必要である。その場合、結果をどの時点でどのような形で公表するかが問題になる。第1の事例は、学問的成果が専門雑誌より先にマスコミに報道され混乱を起こしたフィジシャンズ・ヘルス・スタディを取り上げ、その経緯を紹介する。同じ薬物や同じ治療法に関して世界中で同様の研究が遂行されている。そして、個々の医学研究の結果には相互に矛盾するものがあり得る。そこで正しい結論を得るために総合評価（メタアナリシス）が有用であり、循環器疾患やがんの分野では新しい研究方法として多くの応用事例が公表されている。その際、肯定的な結果のみが公表されやすいという出版バイアスが問題になる。第2の事例として、出版バイアスと、それを避けるための手段でもある臨床試験の事前登録を紹介する。エイズの治療薬についての研究は、その重篤性と緊急性から、従来の研究方法・法的手続きに再検討を迫りつつある。第3の事例では、ある薬剤の認可の経緯を紹介し、主に臨床試験に関する問題提起を行う。

### 1. はじめに

医学に関係する情報の社会的流通と医学研究倫理について若干の問題指摘を行う。そのために三つの事例を取り上げる。第1の事例は、フィジシャンズ・ヘルス・スタディ (Physicians' Health Study) [1] という、最近大きな話題になった、心筋梗塞予防に関する大規模臨床試験であり、研究成果の公表とくにマスコミ報道の問題に触れる。第2の事例は、最近ランセット (Lancet) 誌上に発表された出版バイアスの調査の問題 [2] であり、メタアナリシスと臨床試験の事前登録制度を紹介する。第3の事例は、最近認可されたエイズの新しい治療薬である DDI [3] に関わる問題であり、より一般的に、エイズが与えつつある臨床研究のあり方に対する再検討の動きを紹介する。いずれも、提起する問題について解答を与える意図はないし、また統一かつ明快な解答は不能な問題ばかりである。しかし、情報と倫理の関わりという観点において、統計学の応用に携わる者にとっては、分野によらず深刻かつ興味深い事例と考える。

### 2. フィジシャンズ・ヘルススタディ

1988年の冬に、フィジシャンズ・ヘルス・スタディ (以下 PHS と書く) の結果が新聞報道され、ついでニューズウィーク誌などの雑誌に取り上げられ、大きな反響を呼んだ。微量のアスピリン (2日に一度 355 mg) を飲むことにより心筋梗塞のリスクが半分になるとというのが結果の概要であり、これまでいくつかの先行研究で示されてきたアスピリンの抗血小板凝固作用が

\* 東京大学医学部健康科学・看護学科, 〒113 東京都文京区本郷 7-3-1

† : 本論文は、日本統計学会 60 周年記念事業の一環として、1991 年 11 月 15 日東京で行われたシンポジウムで発表した報告をまとめたものである。

心筋梗塞1次予防において劇的に検証されたのである。最近の研究の中には1日30mgでも薬効は十分であるとするものがあるが、残念ながらわが国における十分なデータはない。

この研究結果は1988年1月27日にわが国の新聞でも報道されたが、アメリカでも同じ日にニューヨーク・タイムスなどの一般新聞で報道された。

この研究は、大規模、長期、前向き、ランダム化、盲検化という条件を満たす1次予防研究であり、被検者は健康な医者集団である。実験は、実験計画法でいう $2 \times 2$ の要因計画で行われ、2万2000人の医者がボランティアとしてこの研究に参加した。全員はランダムに $2 \times 2$ の4群に分けられ、アスピリンとベータカロチンという二つの薬を飲むか飲まないかを割り当てられた。アスピリン服用は冠状動脈疾患すなわちいわゆる心筋梗塞の予防のためであり、ベータカロチンはがんとくに肺癌の予防のためである。薬を飲まない群には、偽薬が与えられ、実際の薬を飲んでいるかどうかは被験者にはわからなくしてある。

長期大規模臨床試験においては通常のことであるが、この試験には、試験の安全性と倫理性を確保するため、データの信頼性及び結果の中間検討を行うモニタリング・ボードが設けられ、これが半年に1度データを見直していた。このボードが研究途中で、アスピリンを飲んだ群と飲まない群の間に心筋梗塞発症に関して圧倒的な違い(服用群が約半分)があることを確認し、1987年12月18日に、研究の中途打ち切りを勧告した。これ以上試験を続行することは非倫理的であるという判断である。

中止の勧告を出したすぐ後に、研究者達は緊急報告として論文を書きあげた。これは直ちに、ニューイングランド・ジャーナル・オブ・メディスン(以下ではNEJMと書く)という臨床医学に関しては最高の権威をもつ医学雑誌に送られた。NEJMは、異例にも2週間で審査を終え、3週間で論文を印刷し、1月28日に発行するという措置をとった。ところでNEJMは、新聞で一度報道された研究は一切雑誌に載せない、ということ編集の原則としている。これはこの原則を定めた当時の編集長の名前をとり、インゲルフィンガー(Ingelfinger)の原則と呼ばれているが、糖尿病に関する大規模研究UGDP(ユニバーシティ・グループ・ダイアビータス・プログラム)の苦い経験を踏まえたものである。すなわち、UGDP研究では、広く使われている経口糖尿病薬が偽薬(プラセボ)に比べ有効ではなく、それどころか死亡のリスクを増やすという研究成果が、論文として丁寧に評価される前に新聞発表され、糖尿病患者の間でパニックを起こしたのである。この研究の計画と成果には現在でも議論が続いているが、いずれにしても「微妙」な問題であり、このとき問題となった経口糖尿病薬は現在でも使われている。この微妙な研究成果が十分な検討や批判の前にニュースとして一般に報道され、アメリカ中の医者に患者の問い合わせが殺到することになった。ところが、医者の方はこの研究成果についてほとんど情報がないまま対処しなければならず、パニックが起きたのである。

この原則に基づいて、NEJMは、雑誌の発行日である1988年1月28日にこのニュース報道を解禁するという約束を報道各社と行った。すなわち、1月21日に刷り上がった印刷論文について、1月28日まで公表をしないという条件で250部をマスコミ各社に送った。そして、1月25日には18万5000部を会員などに発送し、27日、28日までに読者の90%がこれを受け取れるよう手配した。ところが通信社ロイターがその解禁約束を破り、1月28日の前にこの研究成果を世間に報じた。その結果、ニューヨーク・タイムズなどの報道各社はこれに追隨して報道を行ったわけである。

NEJMはロイターに対して、「今後6カ月間ニュースを渡さない」という制裁措置をとった。しかしこの後、2月10日には、この措置を厳しく批判する読者の声がニューヨーク・タイムズに載った。もし12月18日の段階でこのニュースが明らかになっていれば、雑誌が出る1月28日までの間に、何百人かの命が救われたはずだという意見である。この研究の対象として実際

に薬を飲んだ医者が投書してきたのである。

これに対して NEJM は過去の痛い教訓を挙げ、「きっちりとした評価がないまま報道されることを恐れて、この措置をとった」と弁明している。しかし報道各社は、NEJM の措置は偏狭に過ぎるというコメントを行った。NEJM は不十分で不確かな情報を患者が得ることにデメリットがあるという判断をしたわけであるが、これだけ圧倒的な差が出ていれば科学的には問題が無い、それなのに報道を遅らせたのは非倫理的であるという立場である。

学会直前になると、抗腫瘍効果をもつ有望な物質や新しい有効な治療法が発見されたという記事が新聞紙上をにぎわすことが多い。これらの「薬効・効果」の多くは、厳密な臨床研究によってうたかたの如く否定される。しかし、親近者に進行がんの患者をもつ読者がどう記事を受け取るであろうか。われわれは、(私見ではあるが)厳密な科学的評価が十分ではないまま広く使われるに至りさまざまな問題を引き起こした丸山ワクチンという事例をもち、一方で(がんへの薬効は現在ではほとんど否定されている)腫瘍死因子 TNF を妻に使うために輸入を認めてほしいという 1 市民の訴えを一つのきっかけとして対がん 10 年計画が動きだしたという事例も合わせ経験している。医学研究成果に伴うあいまいさとその影響の大きさを考慮したとき、公表の問題をどう考えるべきであろうか。

### 3. 公表されるデータに内在する偏り

まず、医学研究のデータが大変ばらつく例を示す。表 1 は、古典的な抗がん剤 5FU を大腸がん患者に投与したとき、がんが縮んだかどうかを割合(奏効率)で評価したものである。各数値はそれぞれが独立な一つの研究の結果である。非常に効いたとする研究には、奏効率 85% というのもあるが、効果が弱いものには奏効率が 8% というのもある。同じ薬、同じ疾患なのに、

表 1 5FU の大腸がんに対する効果

研究者	患者数	奏効率 (%)
Sharp and Benefeil	13	85
Hall and Good	19	63
Rochlin <i>et al.</i>	47	55
Allaire <i>et al.</i>	17	47
Cornell <i>et al.</i>	13	46
Everly	12	42
Field	37	41
Bell	22	36
Weiss and Jackson	37	35
Ferguson and Humphrey	12	33
Hurley	150	31
Eastern Cooperative Group	48	27
Talley	271	21
Hyman <i>et al.</i>	30	20
Moore <i>et al.</i>	80	19
Ansfield	141	17
Mayo Clinic	358	17
Ellison	87	12
Kennedy	22	9
Knoepp <i>et al.</i>	11	9
Olson and Green	12	8

\* Moertel による

医学研究、臨床研究のデータはこのようにばらつく。どうしてこのようなばらつきが生じるのであろうか。患者層が微妙に違い、5FUの投与方法が微妙に違い、がんが縮んだかどうかの評価が微妙に違い、というのが主な理由であろう。

ところでこのデータには面白い傾向がある。各研究を比べると、患者数が少ない研究ほど成績がよい、すなわち奏効率が高い傾向がある。おそらくこれは、出版バイアスの反映であろう。つまり、研究を行い良い成績がでたら論文を投稿し、うまくいかなかったら投稿しないという傾向のため、また、編集者も良い成績の論文の出版を好む傾向が強いため、そして1施設で行う小規模な研究ほど研究公表の圧力が弱く、かつ世界中を見回したとき実施されている数が多いため、小さな研究ほど実際以上に良い成績が報告されるのであろう。

出版バイアスが実際にどの程度あるかについてはいくつかの論文が公表されている。ランセット誌の論文[2]は、オックスフォード大学の倫理委員会で審査された487の研究を取り上げ、ほぼ半分しか結果が公表されていないこと、さらに、統計的な検定の結果が有意であったものほど発表の割合が増えていることを指摘している。

このような出版バイアスが存在するとすれば、メタアナリシスという新しい研究方法論が危機に陥ることになる。メタアナリシスとは、小さな治療効果や特殊な患者層に対する薬効のように、一つの研究では評価できない問題を、独立な複数の研究を総合することで評価しようとするものである。すなわち、研究成果を何らかの方法で収集し、効果の一様性を検討しながら結果の統計的併合を行い、答えを出そうという方法論である。元来は教育学や社会学の分野で使われてきた方法論であるが、医学分野では、循環器疾患やがんの分野で最近よく使われるようになっていく。

たとえば最近発表された、乳がん治癒切除後の薬物療法の論文[4]においては、全体で133試験、75000人の患者データに基づき、たとえば50歳以上の女性にホルモン剤タモキシフェンを投与したとき、ハザード比で20%の死亡率減少が生じるという結果が導かれている。

治療法の評価と治療指針の設定のため、メタアナリシスは大変有効な手段であるが、公表された論文に基づいて解析を行おうとすると、出版バイアスをどうやって避けるかが問題になる。統計的な解析手法でそれを排除することは(適当なモデルを前提して手法を提案することは可能であり、またそのような提案を行っている研究者も存在するが)健全な態度とは思えない。現在臨床試験に関わっている多くの統計家から提案されている対策は、臨床試験の事前登録である。たとえばがんの分野では、アメリカの国立がん研究所が資金援助を行っている研究のすべてにおいて、プロトコル(研究計画書)が事前登録されている。研究登録を行えば、出版バイアスを避けたメタアナリシスが可能となり、その登録情報が広く行き渡れば、研究協力と現状の把握も同時に可能となり、そして恐らく、むだな研究の重複も避けられる。倫理的には大変望ましいことであろう。

がんの分野では登録された結果をデータベースとして流通させるシステムが既に存在する。フィジシャンズ・データ・キュアリー(PDQ)[5][6]がそれである。われわれは、オンラインのデータベースあるいはコンパクトディスクの形で、これを利用することができる。毎月改訂されるがんに対する標準治療法の情報、患者を登録している約1000の研究のプロトコル、これらの研究に参加している施設と医師名が簡単にわかるシステムである。このシステムを使うことにより、開業医は、最も進んだ研究に患者を登録することができる。このシステムは、がん死亡率減少をめざしたアメリカ国立がん研究所の長期プロジェクトの産物であり、最近ではアクセスサービスも始めているようである。

知識の普及の問題に関し、最近発表された「出版バイアスと臨床研究の普及」という論文[7]で大変面白い過激な提案がなされている。すべての研究は事前登録し、その結果をオンライン

誰でも利用できるようにする。雑誌にはオリジナル論文を載せないようにするというのである。雑誌は多くの研究を比較し、レビューし、読者を教育し、討論の場を提供することを役割とすべきだという提案である。ちなみにこの論文の著者たちは、アメリカ東海岸のがん共同研究 (ECOG) に携わる生物統計学者であり、「統計的検定が研究方法論として考え出されたフィッシャーの時代とは研究環境が異なってしまった」というのがこの論文の書き出しである。

#### 4. エイズ治療薬の波紋

最後にエイズの問題を取り上げよう。新しくエイズの治療薬として認可された DDI という薬は、今までの常識を破り、申請後わずか3カ月、人間に投与されだしてから3年半という近年の最短レコードで認可された。

一般に新しい薬が認可されるには大変複雑なプロセスが必要である。合成と探索から開発が始まり、認可されるまで、通常は約10-15年の年月と80-150億円の費用がかかるといわれている。申請された後に、アメリカの認可当局 FDA が審査に使っている月数の平均値は、1988年には31カ月であった。ところが DDI の場合はこれが3カ月であった。問題の重篤性・緊急性と患者集団の強い圧力がこの短縮の理由である。この研究が行われていることは基礎の段階からマスコミに報道され、インビトロ (試験管内) での非常に希望のもてる成果が報道され、大きな期待と圧力のため、FDA はこの審査を3カ月で済ませたことに加えて、臨床試験研究に入らない患者に対してもこの薬を出すという異例の措置をとった。パラレル・トラックという制度である。

このように、エイズは今までの薬物評価、認可、そして臨床試験の方法論を大きく変えつつある。当然、臨床試験に関わる生物統計家たちもこの変化に関心ではありえない。1990年には、約20人の統計家が、臨床試験研究に与えたエイズの影響を分析し、今後の臨床試験のあり方に対する提言を行っている [8]。そこでは、いかにすれば倫理性を考慮した迅速な試験が可能となるか、対照群としてプラセボが許される研究とはどんなものであるか、どのような状況のもとでは対照群なしの研究が valid であるかなどについて真摯な議論がなされている。(筆頭著者であった David Byar は1991年夏にエイズで死亡した。) 臨床試験の方法論の古くて新しい問題が、改めて真剣に議論される状況になっているのである。

#### 5. おわりに

とくにがんの分野を中心として臨床研究とくに臨床試験の方法論の進歩はわが国でも著しい [6] [9] [10]。盲検化できない研究の場合の電話・ファックスによる患者登録制度、コンピュータを用いる動的割付法、長期試験におけるモニタリング制度、高度な統計解析手法の応用は今や常識である。図1は1992年から開始予定の脳梗塞予防の大規模試験の研究組織である。日

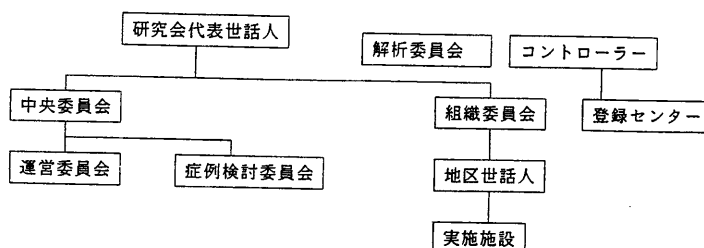


図1 ある臨床試験の研究組織

表2 統計学者の生産

	アメリカ	日本
統計学科	多数	0
生物統計学科	約60(Biometryを含む)	0
統計学会員	15000強	1000強
内生物統計関連	1500	計量生物学会員300
Ph. D. (1980-82)	647	22
内生物統計関連	121	?

(学術会議勧告「統計学の大学院研究教育体制の改善について」1983年11月18日から)

本人には苦手といわれてきた研究の組織化も可能となってきたのである。

このような動きに対し、生物統計家の教育と供給が緊急の課題となりつつある。すなわち、以上提起したいずれの問題においても、生物統計家の寄与が過去になされ、また現在も必要とされているのである。わが国のこの分野の人材不足は危機的状況である。最近の統計によると、アメリカにおいては、Biometrics, Biostatisticsの大学院を持っている大学がそれぞれ30ほど存在する。一方わが国において生物統計を冠した講座はわずか2つである(広島大学原爆医学研究所, 1992年から東京大学医学部健康科学・看護学科(旧保健学科))。表2は、丘本正教授がまとめた日本の現状のデータである。解説は不要であろう。この分野の統計家の育成をどうするか、日本統計学会でも本気で検討をすべきでなかろうか。

## 討 論

### 吉村 功(東京理科大学工学部)

大橋報告にコメントする。医学データをどのように一般の人に知らせるかは非常に重要な問題である。たとえば報告で提起された問題は、一般化すると、第1種の過誤と第2種の過誤をどうバランスさせるかという問題である。しかし一般の人には、そういう物の考え方がない。だから一般の人は、その判断においてパニックを起こしやすい。

インフォームド・コンセントでも、一般の人には、適切に判断する能力がない場合が多い。そこで今必要なのは、そういうことを理解できる人をふやし、皆がパニックを起こさないで、時々刻々ふえる情報を利用し、それ応じて適切な行動をとれるようにすることである。そのための方法を考え、かつ実現することである。われわれとしては、公衆にこういう類の情報を、理解できる形で伝達するシステムを作ることと、それを機能させるよう努力することが必要である。時間はかかるがそれ以外に問題を解決する方法はないのでなかろうか。

### 大橋靖雄

私も最後に残るのは同意の問題であると思う。そのためには、中学校レベルの保健・体育の教育を少し変えることも考えるべきであろう。医療を受けるというのはどういうことか、社会的なコストと自分のベネフィットをバランスさせるというのはどんなことか、治療法の進歩のために臨床的な研究がどういう形で必要か、というようなことをかなり初期から教えた方がよいのでなかろうか。

アメリカの国立がん研究所では、臨床試験に参加する患者のためにすばらしいパンフレットを用意している。そこには、がんとは何か、化学療法とは何か、同意のプロセスとはどういうものかが判りやすく書かれている。それを読むことで、今何が自分の身に起きているかを自分

のペースでゆっくり理解できるようになっている。日本の医療の現場ではそういうことが行われていなかったり、あるいは行おうとすると、基礎知識が不足しているために時間と労力を大変多くかけなければならなかったりする。応用統計家にとっては、このような問題は異次元の問題ではない。むしろ積極的に取り組むべき本質的問題と考えている。

#### 参 考 文 献

- [1] Steering Committee of the Physicians' Health Study (1989), Final report on the Aspirin component of the ongoing Physicians' Health Study, *The New England J. of Medicine*, **321**, 129-135.
- [2] Easterbrook, PJ, Berlin JA, Gopalan, R and Matthews, DR (1991), Publication bias in clinical research, *The Lancet*, **337**, 867-872.
- [3] Hanna, KE ed. (1991), *Biomedical Politics*, National Academy Press.
- [4] Early Breast Cancer Trialists' Collaborative Group (1992), Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy, *The Lancet*, **339**, 1-15, 71-85.
- [5] Hubbard, SM, Henney, JE and Devita, VT (1987), A computer data base for information on cancer treatment, *The New England J. of Medicine*, **316**, 315-318.
- [6] 大橋靖雄 (1991), 癌治療比較試験におけるコンピュータ利用の現状, *癌と化学療法*, **18**, 1403-1408.
- [7] Begg, CB and Berlin, JA (1989), Publication bias and dissemination of clinical research, *J of National Cancer Institute*, **81**, 107-115.
- [8] Byar, DP et al. (1990), Design consideration for AIDS trials, *The New England J. of Medicine*, **323**, 1343-1348.
- [9] 大橋靖雄, 小川一誠, 下山正徳 (1991) 中里博昭, 臨床試験の現状と将来, *Oncologia*, **24**, 57-64.
- [10] 大橋靖雄 (1989), 癌臨床試験の統計的側面と情報支援, 1989年度尿路悪性腫瘍研究会記録