

# 米国における公的研究開発の評価手法

2002年5月

文部科学省科学技術政策研究所

第2研究グループ

Diana Hicks, Peter Kroll, Francis Narin,  
Patrick Thomas, Rosalie Ruegg,  
富澤宏之, 齋藤芳子, 小林信一

本研究は、科学技術政策研究所が米国の CHI Research, Inc. と協力して実施した。著者のうち、Diana Hicks, Peter Kroll, Francis Narin, Patrick Thomas は CHI Research, Inc.に所属し、Rosalie Ruegg は Technology Impact Assessment に所属している。ただし本資料の作成は、科学技術政策研究所の事業として実施した。また、調査研究にあたっては科学技術振興事業団の協力を得た。

Quantitative Methods of Research Evaluation Used  
by the U.S. Federal Government

May 2002

Diana Hicks<sup>†</sup>, Peter Kroll<sup>†</sup>, Francis Narin<sup>†</sup>, Patrick Thomas<sup>†</sup>,  
Rosalie Ruegg<sup>‡</sup>, Hiroyuki Tomizawa, Yoshiko Saitoh,  
and Sinichi Kobayashi  
(<sup>†</sup> CHI Research, Inc., <sup>‡</sup> Technology Impact Assessment)

2nd Theory-Oriented Research Group,  
National Institute of Science and Technology Policy (NISTEP)  
Ministry of Education, Culture, Sports, Science and Technology (MEXT)  
JAPAN

## はじめに

本報告書は、米国で用いられている研究評価のための様々な定量的手法について報告し、日本における定量評価の取り組みの参考とすることを目的としている。とりあげたのは、連邦政府の研究開発プログラム／プロジェクトを対象とした定量的評価手法であり、特に、プログラム／プロジェクトの結果や成果の評価のために用いられた経済学的手法および計量文献学的分析手法に焦点をあてた。

米国の研究評価は、伝統的にピアレビュー、すなわち研究内容を理解できる専門家による評価が中心であり、定量的評価手法の活用が盛んになったのは比較的、最近のことである。その背景には、業績・結果重視の行政管理の導入や、国民への説明責任の重視といった趨勢がある。本報告書でとりあげる研究評価手法は、このような状況のなかで、その適切さを常に問われ、批判にさらされることにより発展してきた。

一方、我が国においても公的資金による研究開発支援の強化に伴い、その効率的な運営と国民的理解の獲得のために研究評価システムの整備が図られつつある。また、日本の行政全体が政策評価に取り組み始めているなかで、科学技術関連行政に関しても、その政策対象ではなく政策自体を評価する試み、すなわち、これまでほとんど行われていない研究開発プログラムや施策の評価も始まっている。定量的手法は、万能ではなく適用できる範囲は限られるものの、これらの評価の取り組みにおいて不可欠なツールである。しかし、経済学的手法はもちろんのこと、計量文献学的分析手法についても経験の蓄積はほとんどなく、取り組みも盛んであるとは言い難い。

評価手法は、どのような対象にも常に適用可能な一般的方法は存在せず、評価対象ごとに慎重に設計されるべきだと言われている。そのため、米国で用いられた評価手法をそのまま我が国に導入するのではなく、各手法の特性およびその手法が用いられた状況を理解した上で、より適切な活用法を探る必要がある。

本報告書の作成にあたっては、単に個々の手法の技術的な解説にとどまらず、米国の定量的評価手法について、それが用いられる文脈も含め、多面的な情報を提示することを目指した。本報告書が、我が国の研究評価の取り組みに有益な示唆を提供できれば幸いである。

# 目 次

はじめに

第1部 概要と政策的含意	1
第1章 研究の目的	3
第2章 報告書第2部の概要	6
第3章 政策的含意	16
第2部 米国における定量的研究開発評価手法	19
第1章 研究開発評価の概観	21
評価の基本的性質	22
評価者と評価対象	28
結論	33
第2章 ATP(先端技術プログラム)で用いられている評価方法	35
評価の折衷的アプローチ	36
ポートフォリオ属性の統計的特徴	37
関係者への調査	37
成果 (outcome) 指標	41
事例研究法	42
統計的・計量経済学的手法	52
未公表の統計的・計量経済学的研究	55
要約と結論	56
第3章 計量文献学	59
序文	59
歴史と妥当性の確認	59
長所と短所	69
方法の要素	81
結論	89
第4章 計量文献学を利用した評価の例	91
A．米国連邦政府のために行った定量的データを含む最近の研究評価の例	95
B．論文間の分析：研究結果 (output) の評価	109
C．特許間の分析：研究成果 (outcome) の評価	143
D．特許・論文間の分析：研究と技術の連関 (linkage) による成果 (outcome) の評価	157
第5章 参考文献	193
附属資料 主要文献の概要	199
用語解説	209

# 第1部

## 概要と政策的含意

富澤 宏之

齋藤 芳子

小林 信一



## 第1章 研究の目的

本書の目的は、我が国において研究開発の評価に取り組む関係者が直面している「適切な定量的評価手法をいかに選択するか」という問題に対する参考資料を提供することである。そのために、米国において連邦政府の研究開発を評価するために実際に用いられた手法、特に定量手法についての包括的な調査を行った。

本書の主たる読者としては、政策策定者や R&D 評価の実務者を想定した。また、科学技術政策や研究開発マネジメントの研究者は、本書によって米国における R&D 評価の手法の全般的な状況を把握することができるであろう。より詳しい調査・分析の手がかりとして、詳細な参考文献リストも付けた。

本書作成の基本的な方針は以下の通りである。

### 1. 対象の範囲

本書が定量的評価手法を題材としたのは、我が国において、定量的評価手法についての理解が不足していると考えたためである。定量的評価手法は、評価の方法論の一部分に過ぎず全面的な依存は危険であるものの、評価対象の特徴や評価の目的に合わせて適切な手法を選択して用いるならば、他の方法では決して代替できない情報源となる。

米国連邦政府の資金によって行われた研究開発プログラム／プロジェクトに対象を絞った理由は、その評価の取り組みが批判にさられるがゆえに手法とその有効性を向上させてきたこと、情報公開が進んでおり比較的偏りの少ない分析が可能であること、そして何よりも我が国の関係者の関心が高いことによる。

米国においては、業績・結果重視の行政管理が導入されるなかで、様々な領域の評価に関して、結果(output)と成果(outcome)を区別し、しかも後者をより重視しようとする傾向がある。科学技術政策の領域でも、最近では研究開発のもたらす成果(outcome)の評価への関心が高まっている。しかし、日本では研究の成果(outcome)の評価だけでなく、結果(output)の評価についての取り組みも盛んでなく、一般の理解が深まっていないと考え、両方についての解説をバランスよく盛り込むこととした。成果(outcome)の評価も、結果(output)の評価に関する経験の蓄積が基礎となることを強調したい。

技術拡散、技術移転の評価については詳しく述べていない。しかし、米国においては目的志向の研究開発の評価が重要になっており、本書でとりあげた研究開発プログラム／プロジェクトの評価のなかには、技術拡散、技術移転の局面も対象とした評価の事例が含まれている。

研究開発のマクロ的な評価、すなわち研究開発の経済成長への貢献についてのマクロ経済学的研究は、世界的に興味を持たれている話題であるが、今回の調査の範囲を超えるテーマである。なお、これらの研究は、本書でとりあげた研究開発プログラム／プロジェクトの評価手法とは直接的な関係はないものの、定量的研究評価の基本的な考え方に示唆を与えることは確かである。

本書でとりあげなかった項目のひとつが、研究者の個人評価のための手法である。研究者評価に対して定量的評価手法を適用することは、必ずしも適当でない場合が多く、本書のような定量的評価手法を中心とする解説書・参考書の範囲を超える問題と考えたためである。

## 2. 具体的な解説

本書の中心は、研究開発評価で用いられている手法(評価のために行われる分析の手法も含む)の具体的な解説である。ただし、個々の評価手法の「使用説明書」でなく、それぞれの手法の基本的な考え方に重点を置いた。また、それぞれの手法の短所と長所、問題点、その手法を活用する際に注意すべき点、基礎となるデータベースなどについても述べた。

ケーススタディやピアレビューについては、具体的な手法を詳しく説明していないが、その重要性、定量的手法との関係や位置づけについて記述した。

## 3. 評価手法をとりまく状況・文脈についての記述

ある評価手法について、その手法が生まれ、あるいは発展した状況を知ることは、その手法を適切に利用するために重要である。そのため、個々の手法が用いられる状況や文脈についての記述を重視した。また、評価対象となる R&D プログラム／プロジェクトや組織・機関、制度について説明し、評価手法の具体的な説明の基礎とした。

さらに、それぞれの手法がおかれた状況だけでなく、第 2 部の第 1 章を中心に、米国における公的研究開発の評価の全般的動向についても紹介した。特に米国では、政治・経済情勢の変化に応じて、公的研究開発を取り巻く状況が大きく変化することがあるので、このような全般的動向の理解は極めて重要性が高い。



このような方針を実施するため、本調査研究では、科学技術政策研究所(NISTEP)と米国の専門家が密接に連絡をとりあい作業を進めるという形態をとった。米国における研究開発活動に実績を有する CHI リサーチ社のスタッフ、ATP(先端技術プログラム)における研究評価に実際に携わってきた Rosalie Ruegg が調査に参加した。NISTEP 側のデマンドを米国の文献に適合するように調整を繰り返しながら、米国側が資料をまとめ、NISTEP 側はその翻訳と日本の文脈での解釈を行った。米国側のスタッフの Diana Hicks は、かつて NISTEP に特別研究員として在籍したこともあり、日本側のデマンドをよく理解し、橋渡し役を果たした。また、調査の実施に際しては、科学技術振興事業団(JST)の協力を得た。

## 第2章 報告書第2部の概要

### 第1章 米国における公的資金による研究開発の評価の概観

米国においては、これまでピアレビューによる評価が優勢だった。しかしながら、最近ではピアレビュー以外の評価も拡大している。

ピアレビュー以外の評価活動の拡大には、以下の理由がある。

第1に、そして最大の理由として挙げられるのは、政治的風土が変化したという仮定である。1991年のOTA(技術評価局)の報告では、研究評価の需要の増加は、予算制約や後援者(議会、納税者)に対する説明責任の拡大、および意思決定をもっと合理的に行なおうとする欲求などに対応するためだったとしている。1993年には、基礎的研究を行うNSF(全米科学財団)でさえ、プログラムの大半を国家的目標に結び付けて、設定したマイルストーンに対する進捗状況を報告するように上院議員から要請され、研究コミュニティは衝撃を受けた。

第2に、実際の研究結果(Output)とは別に研究成果(Outcome)の評価を強調する傾向が強まった点である。研究結果は、出版物や科学的評価などに見られる新知識や科学的優秀さなどの伝統的な研究の産物であるが、研究成果とは社会的・経済的利益、新技術、環境改善など、国がその研究投資から得る成果を意味する。結果の評価はピアレビューで評価可能であるが、成果の評価はピアレビューでは難しい。また、伝統的な論文間の計量文献学的分析手法は、研究成果の評価にはやや不適切であると見られている。

また近年は、研究のための研究は、技術やその他の用途と直接結びついた研究ほど利益はなくなっている。これは研究の特性が変わってきていることを示しており、今日では研究機会が多くなればなるほどそれだけ、例えば、生命科学や情報技術などにおける技術に密接な結びつきができる。*Nature*と*Science*の両誌は特許にもっとも頻繁に引用される雑誌である。特許と論文のあいだの計量文献学的評価手法は、こうして研究成果の評価ツールとして利用できるようになってきた。

さらに、実績(Performance)や成果(Result)への議会からの関心が、GPRA(1993年政府業績・成果法)に示されている。

近年ATPは、米国における評価風土を全面的に変えてしまった。同機関は多くの民間の自営コンサルタントや学界に評価研究を委託し、そしてこれらの研究を公表した。この研究は主として

経済学的なものだった。ATPはそのスタッフとしてエコノミストを雇用し、評価に適用する経済学的方法の開発を推進した。

ところで、連邦プログラムの評価の細分化傾向により、評価活動を包括化することが困難であるが、それは評価の公表が限られていることから生じている。しかしながら、インターネットの利用により、未刊行の評価研究が広範に利用できるようになり、従来より長期間保存しておくことができるため、この細分化の傾向は弱まっている。

## 第2章 ATP(先端技術プログラム)で用いられている評価方法

NIST-ATP(国立標準・技術研究所-先端技術プログラム)は、1990年に当該プログラムが発足した当時から、評価プログラムの開発を開始した。ATPは、民間企業に対する研究開発支援のパートナーシッププログラムであり、米国の産業競争力の強化、米国経済への利益を目的としているため、経済的波及効果が高く、挑戦的でリスクの高い技術への助成を行っている。

そのため、プログラム評価を適正なプログラム管理の要であるとしており、また、1993年の政府業績・成果法(GPRA)と、ATP外部からの恒常的な情報の要求により、評価の必要性に弾みがついた。

ATPの評価プログラムは高い評価を得ている。全米科学アカデミーやその他の提携プログラムが最近個別に行った評価の結論は、「ATPの評価プログラムはその正確さ、範囲および独立性などの点で、米国の他のパートナーシッププログラムをしのぐことは明らかである」としている。評価分野における主要エコノミストであるアーウィン・フェラー博士は、ATP評価プログラムを「他のパートナーシップ活動のモデル」と述べている<sup>1</sup>。政府プログラムにおける浪費、不正行為および腐敗などの摘発を所管する米国商務省監査長官室でさえ、ATPの評価プログラムを「模範的」<sup>2</sup>と評し、賞賛している。

---

<sup>1</sup> National Academies of Science, National Research Council, Board on Science, Technology, and Economic Policy, *Government-Industry Partnerships: The Advanced Technology Program; An Assessment of the Advanced Technology Program* (Washington, D.C.: National Academy Press, 2001).

<sup>2</sup> U.S. Department of Commerce, Office of Inspector General, Report on the ATP.

## ATPにおける評価手法

ATPは業務に最良のツールを使う意気込みで、質と量、両面にわたるさまざまな評価方法を用いてその助成金<sup>3</sup>の効率性を評価した。評価研究が重視する対象は、以下のようにさまざまである。

- ・ ATPのイノベーション・プロセスへの効果
- ・ イノベーターの商業的利益の向上
- ・ 波及効果
- ・ 共同研究の著作
- ・ 評価ツールの改善
- ・ その他、効果の組み合わせ

また、ATPの評価プログラムの主要目標は次の3つである。

- (1) パートナシップ・プログラムの改善のためにさまざまなプログラムの投入、産出および結果などの関係についての理解を深めること
- (2) 何が機能し、何が機能していないかを知るために、個々のプロジェクトの成果を追跡し、測定すること
- (3) プログラムがその目標の達成に向かっているかどうかを評価するために、そのプロジェクトのポートフォリオから得られる大規模な純社会的利益を測定すること。その場合、私的利益、波及的利益、一般的な社会的利益 および ATPに帰すべき利益などに注目する。

ATPは技術開発のための応用研究に資金を提供しているので、評価プログラムはもっと定量的な、成果指向型の経済学的評価方法を推進することができた<sup>4</sup>。ATPが用いた主な評価方法は次の5つである。その他、計量文献学的分析手法や特許引用分析手法なども使用されてきたが、その利用度はこれらの方法より小さい。

---

<sup>3</sup> 資金供与対象プロジェクトの選定（事前評価）はピアレビューで行われているが、本報告書の主題は、資金を提供されたプロジェクトの事後評価である。

<sup>4</sup> 反対に、主に基礎研究に資金を提供している連邦プログラムに適用するには、一般的に経済学的・財政的手法の方がむずかしく、問題が多いことがわかっている。

- (1) プロジェクトのポートフォリオの構成を追跡するための統計的プロファイリング  
助成件数は？ 投入資金額は？ プロジェクトへの参加企業・大学等は？ どの技術分野に資金が提供されたか？ プロジェクトの継続期間は？ 何件のプロジェクトが完了したか？ 実施中のプロジェクトは何件か？ 完了前に中断したプロジェクトは何件か？ その理由は？ 等の質問を行い、データベース化し、分析すること。
- (2) ATP からの資金を受けた者と受けなかった者の調査  
非採択者が各自の研究計画案を継続することができたかどうか、および採択者と非採択者はその後、他の資金供与を受ける能力に差があったかどうか
- (3) プロジェクトの逸話的研究による事例研究(アクネドート事例研究)  
技術開発のストーリー(人間的側面、研究、技術的課題、事後の応用など)を明らかにすることで、プログラムとその機能の態様、研究が公的支援を受ける潜在的価値等を理解・評価する。また明らかにされたプロジェクトの内容を、統計的分析または費用便益分析などの定量的分析に結びつける。
- (4) 費用便益方法、期待値分析および感度分析などによる事例研究  
医療技術等のプログラムを対象として、投資から得られる経済的利益を NPV (Net Present Value: 純現在価値) 等の経済指標で測定する。
- (5) 生産関数、回帰分析、費用指数モデルおよびマクロ経済学モデルなどをはじめとするさまざまな統計的・計量経済学的方法  
高度な技術から得られる消費者利益を推定する新しい方法の開発など。

### 第3章 計量文献学

資金供与額、学生数などの一義的な定量的データと比較して、計量文献学的分析を行うことで、特有の付加価値が得られる。つまり、定量的データからは総計を出すことはできるが、計量文献学的分析データからはさらに引用データが得られ、かつ引用データにはさまざまな用途がある。評価に役に立つ関連データには次の3つの種類がある。

- ・ 論文と論文
- ・ 特許と特許
- ・ 特許と論文

米国では、計量文献学的分析手法はよく利用されているものの、引用を取り扱う上での方法論について、理解不足から批判が起こる場合がある。そのため、評価に関連した引用データの長所と短所等を、表 A、表 B にまとめる。

## 表 A 計量文献学指標の問題点とそれを最小化する方法

Martin および Irvine の古典的表 (*Research Policy*, 12, p.76, 1983)の更新

<ul style="list-style-type: none"> <li>• <b>問題の分類</b> 問題点とされているもの</li> </ul>	それは本当に問題なのか? どうしたら問題を最小にとどめ、または調査の適正な実施によって排除できるか?
<ul style="list-style-type: none"> <li>• <b>論文発表の慣行</b></li> </ul>	
<ul style="list-style-type: none"> <li>1) 論文発表の「くせ」は機関、分野および国によって異なる</li> <li>2) 作為的操作</li> <li>3) 品質を示さない</li> </ul>	<p>属性の一致するグループの比較や、機関、分野または国レベルでの標準化は可能である。論文発表が二次的事業である機関を評価するのに計量文献学だけを用いてはいいけない。</p> <p>計量文献学指標はその他いくつかの評価手法より作為的操作がむずかしい。引用率も分析してみると作為的操作の方法の中には非生産的なものもある。例えば、短い論文をたくさん作成すればするほど、1論文当たりの平均引用件数や被引用度の高い論文の点数がそれだけ少なくなる傾向がある。その他の想定される不正工作の中には合理的である場合もあるが、それは選択の影響、例えば、信望が低い論文や論文発表がむずかしく、影響度が低い分野が評価に反映される場合にかざられる。</p> <p>引用にもとづいた尺度で、インパクトを示すものを使用すること。</p>
<ul style="list-style-type: none"> <li>• <b>参照と引用行動</b></li> </ul>	
<ul style="list-style-type: none"> <li>1) 下位分野により異なる</li> <li>2) 論文貢献の種類(総説、方法論、実験、理論)によって異なる</li> <li>3) 重要な引用</li> <li>4) 自己引用</li> <li>5) 被引用率は論文の寿命によって変化する</li> <li>6) 特別な状況に関しては柔軟性に欠ける</li> <li>7) 時間のずれ</li> </ul>	<p>原データの集計を使用して照合グループ、例えば、1専門分野内のグループを比較することができる。大半の評価には専門分野間の比較を可能にする標準化した引用尺度が必要である。</p> <p>照合グループの比較ができる。引用をインパクト指標とみなすこと。また必ずしもすべての総説、方法論の引用度が高いというわけではなく、すべての理論研究が引用されないというわけでもないことに留意すること。</p> <p>引用を品質または重要性よりむしろインパクト指標とみなすこと。</p> <p>生産性の高い科学者だけが高率の自己引用を達成できる。(理由は自己引用をたくさん作成するには論文もたくさん作成しなければならぬから)。したがって、自己引用は評価結果を変えない傾向がある。出所が明確な場合は自己引用を除外することも可能である。</p> <p>論文への引用は3~4年でピークに達する。発表後最初の4年間などで各論文へへの引用を集計すれば各論文が引用取得に同等な機会を与えられることになる。その影響が後日に明確になる「ダークホース的な」研究の影響は計量文献学や他のいかなる短期的評価によっても拾い出すことはできないだろう。</p> <p>著者や機関は特殊な状況、すなわち、同等の成果をあげている他の者と比較して各自の論文発表や引用集計をやりにくくなる状況に置かれるかもしれない。低得点はつねに何かに問題があることを示す兆候と見なすべきである。その理由を特定するためには、情報通の同僚に相談することができる。</p> <p>評価を急がなければならない場合は、ジャーナルの影響力などの質を示す尺度を引用集計に替えて用いることができる。</p>

---

• **科学引用索引の技術的制約**

- |                      |  |
|----------------------|--|
| 1) 綴りの誤りや作業上の誤り      | 高度なアルゴリズムを用いて参照の対象となる論文と検索用文字列を照合する。   |
| 2) 筆頭著者のみの引用         | 包括的な著作目録を作成し、その中から論文への引用を探すことから始めなければならない。   |
| 3) ジャーナルのみをカバー       | 書籍がない場合は計量文献学の社会科学への適用は制約される可能性もある(もつとも、こうした傾向は変わりつつあるかもしれない)と Hicks は見ている(Scientometrics, 1999)。計量文献学をソフトウェア開発に適用するためには会議発表論文の包括的な採録が必要になる。 |
| 4) ジャーナルのカバーが不十分     | 多くの分野、とくに基礎研究では問題ない。科学に密接に関連しない技術的分野、例えば、生産や製造などに関してはとくに問題あり。  |
| 5) ジャーナルのカバーが徐々に変化する | これが問題となる場合には、ジャーナルの固定セット (fixed subset) を用いることができる。  |
-



表 B 評価に用いる計量文献学主要の長所

計量文献学の長所	
重要性	
<ul style="list-style-type: none"> <li>• 方法論上の長所</li> </ul>	
1) 定量的性質	操作しにくい。さらにチェックするためには「ソフト」ではない短所になりかねない問題に注目することになる。精度と一貫性。
2) 標準化の可能性	それらの周辺への影響は、中心からは見えないかもしれないが、しばしば標準指標で見える場合がある。
3) 直接性	明確で、簡潔で、非常に単純な集計。論文作成上は何の前提もかかされていない。前提があるとすればそれは引用集計の意義に関するもので、評価の当事者全員に既知の事項である。
4) 論文発表と引用にもとづく	科学的著作やコミュニケーションには出版が需要なので、出版から生ずる尺度は評価のためにだけ作成された情報より影響しにくい。
5) 時間の遅れ	引用は出版後 2～3 年でピークに達するので、引用情報は助成金授与時より 5～6 年だけ遅れる可能性がある。しかし技術的適用に関する遅れよりはるかに少ない。
6) 測定が優れている	制御不能になるまでエスカレートせず、たくさんプロジェクト、機関などを検証することができる。各省庁から地方にいたるまでたくさん集計レベルで同じ指標を使用することができる。
7) 成果に利害関係をもたない者が構成する分析であること	客観性が高いこと、党派の影響、同窓生の縁故関係が少ないこと。
8) 控えめ	評価対象に時間や資源を要求しない。したがって、評価の全部費用を比較すると計量文献学は比較的経済的である。
<ul style="list-style-type: none"> <li>• SCI (Science Citation Index) の技術的利点</li> </ul>	
1) 住所はすべて索引化されている。	共同研究が非常に拡大しているため、今日では論文をその他のデータベースに記載されている単一機関のみに割当てることは、その最初の著者に割当てるのと同様に適切なやり方ではない。
2) すべての科学分野をカバーしている	広範に比較できるデータベースで分野間の比較ができるようにするには標準化された指標を作り出せばよい。
3) 上位ジャーナルをカバー	科学論文は非常に広範でこれを包括的にカバーできるものは何もない。引用度が最も高いジャーナル 5,000 誌ないし 7,000 誌をカバーするという選択はカバーすべき主題の選択方法としては賢明である。
4) 引用を含む	引用分析を可能にする。

表 A では、引用分析に伴う共通の問題に対して、当該問題を最小にとどめるか排除する方法論的戦略を示している。表 B では、計量文献学的分析の長所を、各要素が評価の文脈の中でなぜ重要なのかについての説明とともに示している。

特許間の計量文献学的分析手法の長所と短所については、論文間の場合と同じ詳しさでは論じられないが、類似の問題が生じるし、そして大部分は類似の回答を適用できる。特に表 A の「論文発表の慣行」に基づいて記載した3つの問題は、すべて特許分析に適用されるが、解決策も同様である。「参照と引用行動」についての1、4、5および7については、解決策は7のタイムラグを除いて同じである。特許には、すぐに入手できるジャーナルの影響力など、品質の代替的尺度はない。

特許・論文間の計量文献学的分析では、次の点を理解することが重要である。

- (1) 米国特許での引用の意味
- (2) 米国特許制度における情報開示の誘因
- (3) 特許・論文間の参照項目掲載に関する解釈

(1)については、米国および欧州の特許を比較すると明らかである。欧州の特許はそれほど科学論文を参照として記載していない。米国では特許出願者は出願書類とともに先行技術の参照リストを提出するが、欧州では出願者はそうすることを義務付けられておらず、先行技術の検索はすべて特許事務所の職員が行う。したがって、米国特許の参照リスト作成には、完全な技術的内容を熟知している出願者が貢献しているが、欧州では参照項目リストは、特許制度専門の特許事務所職員が作成する。

(2)については、米国の開示規則は出願者に知っているかぎりの先行技術参照の開示を義務付けており、そうしない場合は特許不許可の根拠になることがある。欧州に比べ米国では特許訴訟のリスクが大きくなるため、参照の掲載がますます必要になっている。

(3)については、特許・論文間の引用指標の解釈は複雑である。複雑な社会的過程や特許事務所規則が特許記述の根底にあるからである。特許・論文間の参照記載件数が増加している事実の解釈はさらに複雑である。社会的過程や特許事務所規則が、科学と技術のあいだの変化よりも指標に作用することがあり得る。

しかしながら特許・論文間の引用指標は特別な長所をもった貴重なツールであり、これによりさまざまな科学・技術関係の定量的比較ができる。

#### **第4章 計量文献学を利用した評価の例**

本章の目的は、これまでに実施された計量文献学的分析作業範囲の説明や、各種方法が利用できるさまざまな手段、手法を使って取り組みが可能である、さまざまな問題を提供することにある。

本章では、計量文献学的分析手法を利用し、政府機関がそのプログラムの検証のために委託した研究の要約を紹介している。4部構成となっており、第1部では計量文献学的分析手法を含む最近の評価研究を説明している。さらに、ポーズマン他による最近の定量的事例研究の説明も含まれている(研究価値マッピング・プロジェクト)。

次の3部では、CHIが米国政府のために行ってきた論文と論文、特許と特許、特許と論文間の計量文献学的分析手法を紹介している。

#### **第5章 参考文献**

第5章には研究評価方法論に関する論文のリストを掲げた。とくに重要なものについては附属資料にその内容を簡潔に整理して示してある。なお、この部分は英文のまま収録した。

### 第3章 政策的含意

#### 1. 評価手法分類

本報告書に取り上げた主な評価手法を分類して表 C に示す。いずれも事後評価の手法であり、成果(outcome)指向のものを中心としている。これらの手法は相互に背反するものではなく、組み合わせて用いられることが多い。

表 C 本報告書に取り上げた評価手法の分類

手法		説明
ピアレビュー	パネル評価	専門家による評価、定性的 output 評価に適、outcome 評価には不適
アンケート、インタビュー		調査対象者の個人的見解に基づく、半定量的
トレース		過去に遡って追跡、半定量的 情報量豊富、一般化できない
ケーススタディ	アネクドット的方法	エピソードの集積、半定量的 状況理解に有効、時間がかかる、一般化できない
	費用便益分析	半定量的 応用研究の予備的評価に適、基礎研究の評価は不可、一般化できない
ベンチマーク		半定量的、長短や傾向を明らかにできる 研究機関の方針決定などに適、包括的評価には不適
統計的方法	プロファイリング	定量的評価の基礎データ集積
計量経済学的方法	費用指数モデル	定量的、モデルにもとづく予測の形をとることが多い 応用研究に対する資源配分のための評価に適、基礎研究の評価は不可
マクロ経済モデル		定量的 経済全体と研究活動とは規模がかけ離れているので適用条件が限定される
計量文献学的手法	論文論文引用分析	定量的、包括的 人文社会科学系の評価としては難あり
	特許特許引用分析	定量的 国際比較に難あり
	特許論文引用分析	定量的、outcome 評価に優れる 米国特許に関しては有効
	その他	Web ページの引用分析：検索エンジンの不安定性に難

## 2. 提言

研究評価を本格的に導入するにあたって米国の評価事例から学ぶべき事項や日本が留意すべき事項を以下に指摘する。

### **評価をポリシーサイクルの中に位置付ける**

研究評価手法の発展のためには、その成果を政策立案当局、政策運営当局が真摯に受け止め、それを活用する態勢を構築することが必須である。形式的にのみそれを行い、実質的な活用をしないのであれば(評価をしたことのアリバイとして用いるなど)、研究評価手法はそのようなものとしてしか発展しないであろう。研究評価を政策形成、戦略形成に結びつける努力が必要である。

### **まずは評価を試行してみる**

さまざまな評価手法を試行し、経験を蓄積することによって、評価の手法が発展する。とくに海外で成功した評価方法がそのまま日本に通用するとは限らない。実験的な評価を通して、日本の文化や制度との適合を図り、評価手法の改良を進めるべきである。

### **万能な評価手法はない**

評価手法は、対象プログラムの目的、特性などに配慮して、巧妙に開発される必要がある。何にでも適用できるレディメイドな手法はない。また、使い方を誤ると評価プロセスを誤った方向へ導く可能性があるばかりでなく、研究評価に対する信頼を損ねるであろう。

### **定量的手法は成果・インパクトの評価に向く**

米国における研究の事後評価はピアレビュー(パネルレビュー)が主流である。しかし、とくに成果(outcome)やインパクトの評価に関しては、定量的評価のウェイトが高まっている。

定量的評価は成果・インパクトの評価に向いているが、恒常的な研究資金による課題や機関の評価には不適切である。

### **各定量的手法のメリット・デメリットを理解する**

経済学的手法、財政学的手法は、基礎的研究の評価には適さない。

マクロ経済学的手法は、マクロ経済の規模と研究開発活動の規模が桁違いであるため、正確性に欠ける。何らかの工夫が必要となる。

費用便益分析などの経済学的手法は、事例研究の有力な手法となる。ただし、事例研究を一般化することには限界があることに留意すべきである。

### **恒常的にプロファイリングを行う**

評価の基礎データとなるプロジェクトやプログラムの基本情報を恒常的に収集、整理しておく。またモニタリングのシステムも整備しておく必要がある。

### **評価の経験を公開・蓄積する**

研究評価は分散的に実施される傾向があるが、評価手法の発展のためには、それらの評価手法の適用経験を公開し、蓄積していくことが必要である。インターネットによる公開は有力な手段となる。

### **評価の方法論を育てる**

研究評価方法論の発展のために、国内評価関連機関(資金供与機関、評価研究委託機関、評価研究受託機関、評価研究機関)、評価研究者による「評価手法の研究・評価」のための恒常的プラットフォーム(研究会等)を創設するべきである。

### **評価に関する知識を普及させる**

評価そのものが日本ではまだ十分に理解されていない。評価に関する用語の統一や基礎知識の普及活動が必要である。

## 第 2 部

# 米国における定量的研究開発評価手法





## 第1章 研究開発評価の概観

Diana Hicks CHIリサーチ社

連邦レベルのプロジェクトまたはプログラムの事後評価に関する米国内の文献にはいくつかのテーマが見られる。ピアレビューによる方法は、ワシントンでは定量的方法よりはるかに普及しており、また全米アカデミー(National Academy of Science, National Academy of Engineering, Institute of Medicine and National Research Council)には強力な支持者や関係者がいる。一方で定量的評価方法はそれほど利用されていないと言われているが、こうした判断は欧州事情との比較によるものである。欧州では数十年にわたって評価が当たり前のように行われてきた。米国では定量的評価方法が利用されていない傾向が問題にされ、その理由もあげられ、さらにピアレビューと定量的方法を併用すべきだとの考えも提示されている。米国の学界では評価方法が細分化する傾向も強い。政府省庁は自分たちが委託した調査について必ずしもすべてを知らないことが多々あり、レビュー論文は著者が直接関与した仕事に集中する傾向がある。最終的には、評価結果と評価方法が論争の種になることが多い。評価結果をきっかけにして資源配分をめぐる政策論争が起きると、米国の政策決定システムが適度に開かれているために多くの当事者がこれらの論争に加わる。さらに多くの分析は外部委託先(コンサルティング会社や学界)が行っている。彼らは契約を求めて競争し、自身の方法を押し通すあまり他人の方法を顧みない傾向がある。

本章ではこれらのテーマを論じ、それに沿った文献を紹介する。こうした状況を前提にした場合、本章および報告書全体は、何を明らかにすべきかという著者の立場で決まってくる。CHIリサーチ社は20年にわたって米国連邦政府のさまざまな機関のために計量文献学的評価を行ってきた。このため、我々には刊行されている文献を越える眼識があるが、それでも状況が細かく分断されているため分かりやすい全体像を提示することにも限度がある。本報告は米国で刊行された報告の大半に沿って、我々の仕事に焦点を当てるものである。我々は計量文献学的評価を行っているため、その長所を十分に示す根拠を提示したり、かなり高度なレベルで手法を説明したりできる。こうした内容は米国における研究評価に関する文献に容易には見出せない。

## 評価の基本的性質

この議論のために、評価の主体者は誰かという観点から評価を内部と外部に分類する。内部評価とは科学コミュニティがすべてを行う調査であって、専門家を政府機関内部から選抜するのか、または外部の専門家をを用いるのかについては問わない。内部評価はピアレビューや、プログラムの成果を本に編集するといった(個人的見解による)アネクドットの定量的方法に依存する場合がほとんどである。外部評価は経済学、計量文献学または事例研究手法のようなある種の社会科学的専門知識を使うものであり、定量的方法を用いる場合が非常に多い。外部の専門知識は通常、政府機関の外部にある。ただし、外部評価で用いる専門知識は科学コミュニティの外部にあるが、それが必ずしも政府機関の外部にある必要はない。例えば、ATP は評価に際して部内の経済学者を利用しており、コンサルタントを雇えるのは成果報告を編纂する場合だけである。このように分類は完全に明確ではないが、それでも大半の場合に当てはまる。

米国では内部評価手法、とくにピアレビューが頻繁に用いられる。1980年代に Logsdon & Rubin (1988) が 10 の政府機関で研究の管理と評価を担当する 44 名にインタビューを行い、接触したほとんどすべての機関でピアレビューを広範に利用していることを報告している。その後二人の調査は繰り返されていないようだが、1980年代後半以降事情が変わった兆しも見えない。Logsdon & Rubin の指摘によると、こうした状況は全米アカデミーが評価方法を検討した 1982 年の報告書の勧告に沿っているとしている。全米アカデミーはしばしばピアレビュー実施の委託を受けていた。Kostoff (1993) によれば、非定量的事例研究や、最近の政府機関の成功を示す「ナゲット(訳注:金塊から転じて成功体験のこと)情報」を報告するアネクドット的评价も政府機関により広範に利用されているという。

外部評価手法に関しては、半定量的事例研究の中でも、ある一つの形態が米国において顕著な歴史をもっている。これは一種の追跡(トレース)手法で、まずイノベーションを選び出し、それに先行する研究を特定してゆく。この手法の最初の事例は、国防総省のヒンドサイト(=過去の洞察)プロジェクト(1969)だった。このプロジェクトでは武器システム研究の背景を調べ、高度な武器システムにつながるアイデアが生まれたのは研究者(科学者と技術者)が応用技術者の抱える問題を詳しく知った時だったという事実を突き止めた。これに続いて、全米科学財団(NSF)がいくつ

かのイノベーションを過去に遡って追跡する「TRACES」研究(IITRI、1968)を実施したところ、その時期はヒンドサイト研究よりずっと前に遡ることになり、結果として基礎研究から大きな利益が得られていることがわかった。何年にもわたって新たな研究が続けられた。その例として「TRACES II」(Battelle, 1973)、ガン医療進歩の支援基盤の追跡、そして最近の例では第4章に述べる工学的イノベーションへのNSFの貢献の追跡(Roessner et al., 1997, 1998b)などがあげられる。

こうした研究(トレース手法)の長所と短所は十分理解されている。Georghiou & Roessner(2000)の報告は、基礎研究と応用研究の相対的貢献度合、研究支援の制度的な背景やその源泉などについての豊富な情報がこの方法によって得られるとしている。しかし、個々の研究の意義に関するたいの判断については「追跡」を行う必要がある。技術における先行事例は非常に複雑なため、追跡作業は「有意義」と判断される先駆的研究に集中し、したがってより結びつきが弱いものや袋小路になっている研究は排除されて、間接効果が無視されることになる。またこの方法は費用がかかるために、調査対象はほんの一握りの無作為抽出もされていないイノベーションに限定され、それゆえ結果も一般化できない。この方法はプログラムやプロジェクトの結果を評価するには適さないのである(この方法は過去を追跡するのであって未来を予測できない)。翻って政府機関に裏付け資料を提供するのに適しているのは、この方法が技術進歩に対する基礎科学の重要性を実証しているからである。

外部評価で用いる定量的方法には経済学的分析と計量文献学による分析の2つの大きなタイプがある。これらは本報告書の主題になっており、米国においてこれらの手法を使って何を達成してきたか、こうした技法の長所と短所は何かについて説明される。第2章では国立標準・技術研究所(NIST)の先端技術プログラム(ATP)の評価に用いられた経済学的方法を検討する。ATPはその評価の委託を通じてこれらの手法を最高レベルに高めた。計量文献学は第3章と第4章の主題である。第3章では計量文献学を用いて評価を行うために必要な方法の基礎を説明し、第4章では評価の概要をまとめてある。これらの多くは過去20年にわたってCHIが連邦政府のために行ってきた計量文献学的評価である。ただし、CHI以外による外部評価も4件含まれている。前述のRoessnerらがNSFのために行った追跡、Bozemanらがエネルギー省のために行った大規模な事例研究(研究価値マッピングプロジェクト)、海軍のためにKostoffが行った対話型データマイニングと計量文献学による分析、および全米アカデミーのCOSEPUP(The Committee on

Science, Engineering, and Public Policy) が行った方法論的実験としての国際ベンチマーキングである。これらはすべてが計量文献学の要素を含んでいるが、むしろ定性的評価の例である<sup>5</sup>。

ピアレビュー以外の方法、すなわち外部評価が米国であまり用いられていないという事実は研究評価に関する論評全般に見られる主題である。Kostoff(1993)の指摘によれば、研究の行為や影響および便益の評価のために開発された方法を報告する論文には重要なものが多くあるという。比較的少数ではあるが、こうした手法の一部を現在連邦政府諸機関が実施中である。その中で結果が公表されているものはごく一部であり、かつ意思決定者が取り入れているものはさらに少ない。Kostoff はこれらの方法を利用頻度別に次のように整理している。すなわち、ピアレビュー、非定量的事例研究、アネクドト的方法、定量的方法である。本報告では単なる提案や理念型だけのものは省略し、その代わりこれまでに実際に利用されてきた方法を議論する。その結果、取り上げるべき資料の量は大幅に減っている。

使用頻度の低い方法、つまり定量的方法に注目すると資料はさらに少なくなる。とは言え、取り上げるべきものはかなりある。連邦政府のファンディング機関が過去数十年にわたって経済学的調査、計量文献学的調査いずれについても委託を行ってきたからである。技術評価局(OTA)は1985年に研究評価を論評し、経済学的方法を用いたNASA(米航空宇宙局)の分析例をいくつか報告した。1980年代にはLogsdon & Rubinが、商務省標準局(NBS)がマイクロ経済学的手法を用いて研究成果を検討したと報告した。興味深いことに今日NBSの後身機関である国立標準・技術研究所(NIST)は研究評価の経済学的方法を開発して最高水準に高めている(第2章参照)。

計量文献学もまた米国ではあまり利用されていない。OTAは1985年の報告書において計量文献学的手法を広く取り上げた。国立衛生研究所(NIH)でのCHIの当時の仕事について、OTA、Logsdon & Rubin およびその他が詳細に報告している。これらの調査については本報告書で概要を述べる。Logsdon & Rubin はまた、計量文献学の広範な利用例を見つけ、検証の対象になった政府機関の半数は何らかの形で計量文献学を用いていたという(18件のうち9件)。無論ピアレビューの例は17件と、もっと多かった(Logsdon & Rubin, 1988, 表2参照)。1991年の報告書第2弾において、OTAは方法論がほとんど進展していないことに気付き、利用が少ないという問題や研究をどのように構成すべきかなどの問題に焦点を当てている。これら2つの調査の結論はい

---

<sup>5</sup> COSEPUP 研究は「内部」、その他は「外部」に分類される。

ずれも、定量的方法が米国ではあまり利用されていないということであった。この判断が手法の普及している欧州との比較によるものであることは明らかである。1985年のOTA報告書は英国における計量文献学を論じており、また1991年報告書に編纂されている参考文献には欧州のものが多く含まれている(Averch, 1993 a)。

近年、連邦レベルでの研究評価が増加している。Melkers & Roessner(1997)は、連邦、州いずれのレベルでも評価活動の復活を指摘している。CHIの記録でもこの指摘を裏付けることができ、調査委託がOTA報告書以来大幅に増加していると述べられている。図1-1はCHIリサーチ社が1977年-84年、1985年-92年および1993年-2000年の期間に連邦政府機関のために行った調査の件数を示したものである。OTAの報告書の時期も図中に示した。ピアレビュー以外の評価活動が拡大した証拠は他にもある。すなわち、第4章で要約したCHI以外の調査は1990年代後半に行われている。さらに、研究評価に関する歴史的事例研究を行っているコンサルタント会社Abtアソシエーツの事業記録は、全米科学財団(NSF)のために行った研究を掲載した同社のウェブサイトから入手できる。それによると1991年までは評価方法および大学における論文発表の定量的分析やその他科学政策関連の作業に関して、いくつかの論評が行われていた。1990年代後半には調査が5件あり、その標題はすべて「…の評価」となっている。このようにピアレビュー以外の活動が明白な拡大を見せたことは、いくつかの仮説によって説明できる。

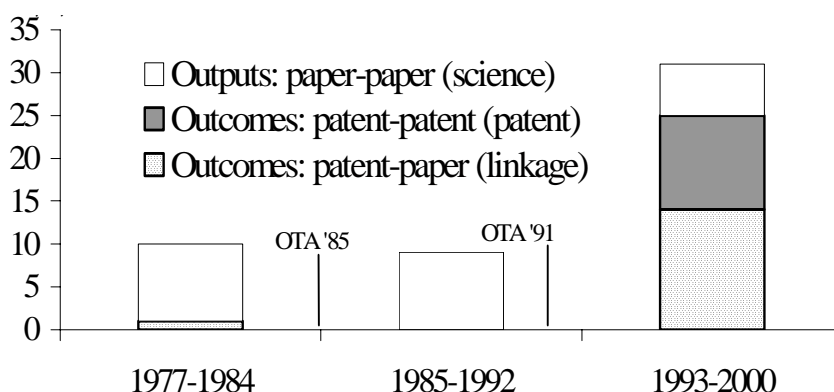


図 1-1 Number of evaluations conducted by CHI for the US federal government

第一の、そして最も支持されている仮説は政治的風土が変化したというものである(Melkers & Roessner, 1997)。当初これは、海外においてもっと顕著だった。しかし1991年のOTAの結論で

は研究評価の需要が米国でさえ増加しており、これは予算の制約、スポンサーに対する説明責任 (accountability) の拡大、および意思決定の合理化への欲求などに対応するためだったとしている。OTA は米国の政府機関にはたいてい評価部門が設置されていたと述べた(1991 年 OTA 報告書、p.252-253)。これらの傾向は 10 年の間に強まったものである。1995 年に Cozzens は、連邦予算(研究資金の資金源)の自由裁量的部分への圧力が強まっており、連邦支出を支配下におこうとする議会固有の動きがあるとの所見を述べた。Cozzens は OTA が指摘している説明責任の拡大というテーマも取り上げ、「さまざまなプログラムが米国民に与える利益を検討しようという真剣な態度が新たに見られるようになった」と指摘している。Barbara Milulski 上院議員は 1993 年、全米科学財団にそのプログラムの大半を国家的目標に結び付け、目標に邁進するための道標 (milestones) を設け、そしてその進捗状況を報告するよう要請して研究コミュニティに衝撃を与えた。Cozzens はこれを冷戦後の時代における「科学と社会の契約」の再考を促すものと見ている (Cozzens, 1995b, p.352)。これらの今まで以上に厳しい条件は、まさに 10 年前の欧州でピアレビュー以外の評価手法を増加させたものと同じである。

第2に、実際の「結果」(output)ではなく研究「成果」(outcome)の評価を強調する傾向の強まりが挙げられる。研究「結果」は研究の伝統的産物で、出版物や科学的名声などに現われる新しい知識や科学上の卓越した成果などを指す。「成果(outcome)」という用語は社会的・経済的利益、新技術、環境改善など、国がその研究投資から得るものを意味する。「結果」の評価はピアレビューの得意とするところであったが、「成果」を重視する気運はユーザーの意向を優先したためピアレビューを主流から外した。同様に伝統的な論文間の計量文献学も研究「成果」の評価にはやや不適切であると見られている (Georghiou & Roessner(2000)が Cozzens et al.(1994)の報告を引用している)。

近年「研究のための研究」は、技術やその他の用途と直接結びついた研究に比べてかなり関心が低くなった。これは研究の性質が幾分変化していることを示している。例えば今日の生命科学や情報技術などにおいては研究するほどに技術と密接な結び付きができていく。Narin は論文を引用する特許のパタンにもとづいて、科学と技術が相互に混ざりあいつつあると論じた (Narin & Noma, 1985)。Stokes は、科学政策は基礎と応用の区分を超える必要があるとしている。今日の仕事の非常に多くは、パストゥールの仕事のように基礎的でもあり、応用的でもあるからである (Stokes, 1997)。CHI は、特許に引用された論文を検証し、引用された著作が基礎科学の雑誌に

掲載されており、また科学論文にもよく引用されていることから Stokes の議論を裏付けた。「Nature」と「Science」両誌は特許にもっとも頻繁に引用される雑誌なのである(Hicks et al., 2000)。品質の高い研究の中に、成果と緊密に結びついたものが増加しているため、そうした結びつきを例証するのは容易になっている。これは評価に対してかなり直接的な影響をもたらし、米国特許による論文引用率の増加は研究と技術との連関をより明確に示すようになってきた。特許と論文の間の計量文献学はこうして成果に関する評価ツールとして利用できるようになり、またこれは1990年代における連邦政府のための CHI の主要な事業の一部になっている(図1参照)。

成果を強調する考えは議会によって強く主張された。実績や最終的成果への議会の関心は GPRA、すなわち、1993年の「政府業績・成果法」にも示されている。Cozzens は GPRA と研究評価との間の関係を分析してきた。この法律は連邦政府機関に、戦略的計画を作成し、毎年実績指標を用いて報告し、戦略的計画や実績評価作業においてプログラム評価を利用することを要請している。研究機関をはじめとしていかなる政府機関もこれを免れず、まもなく完全な実施を行うことになっている(Cozzens, 1995 b)。GPRA が結果(output)と成果(outcome)の間の区別を重要視し、成果を主要関心事としていると Cozzens は書いている。GPRA は各機関に目標の設定とその進捗状況の測定を要求しているため、研究にとって問題をはらんでいる。これは研究活動(助成金の供与件数、助成される研究学生の数など)の役には立つが、政府機関の管轄外である研究の利用または応用のためにはならない(Cozzens, 1995a)。GPRA は他方でピアレビュー以外の評価への関心を高める役割を果たしてきた。Cozzens(1995b)はこう記している：

連邦研究プログラムの評価の大半は記述的であり、GPRA が要求している実績の数量化からほど遠い。数々の評価のための定量的ツールが論文としてまとめられてきたが、実際に用いられているものは皆無に近い。こうして研究プログラムと研究機関は GPRA に対応するために、これまで避けてきた選択肢の中からいずれかを選び出すか、または新しい方法を編み出す、という課題に直面している(p.354)。

GPRA が数量化を要求した結果、ワシントンでは評価方法をめぐって再び議論が起こった。そしてこの議論の裏にはピアレビューの価値をひそかに蒸し返そうとする動きが見え隠れしていた(COSEPUP, 1999)。Cozzens(1995b)は4つの政府機関による先導的プロジェクトによって、数量

化を避けようとする政府機関と測定可能な指標を要求する行政予算局(OMB)とのあいだの緊張関係の図式の中で受入れられる指標が開発されるとみている。

ピアレビュー以外の評価方法の利用を拡大しようという別の要因も働いているかもしれない。助成金交付プログラム(グラント・プログラム)の性格が複雑化していることである。成果重視は科学と産業の結びつきに対する政府の関心を引き起こした。こうした傾向は、技術移転強化の目標の下で国立研究所に産業界と協力する枠組みを提供する協力研究・開発協定(CRADA)などのような、産業界との連携を強めようとする政策とプログラムにつながった。産業界との連携を行っている研究センターに対する資金提供も顕著になっていった。

1990年代後半に行われた評価はこの傾向を反映している。Abt アソシエーツが行った前述の評価はこの新しい複雑なプログラム(SBIR:小企業イノベーション研究、設計・製造・産業イノベーションプログラム、科学技術センター・プログラム、ERC:工学研究センター)に取り組んでいる。こうしたメカニズムは伝統的な科学の分野を超えて優れた研究を波及させることを目標に掲げており、ピアレビューから得られる以上に包括的で仕組みのしっかりした外部評価が必要なのは当然である。伝統的な論文間の計量文献学だけでは不十分になりがちだということから、CHIの業務に変化が起こった。1980年代にはCHIの論文間または科学に関する計量文献学は孤立する傾向があった。90年代後半になって評価の共同実施によりCHIの論文間の分析は事例研究方法と結びついた。こうした手法の複合化への動きは、より複雑なプログラムの評価にはより複雑な方法が必要であることを示唆している。

### **評価者と評価対象**

米国における研究評価は細分化によって特徴付けられる。Melkers & Roessner(1997)は次のように述べている:

「工業国の中で米国の政治制度は際立っている。米国では権限が比較的細分化されていること、多様なアクセスポイントにおける組織化された利害への対応が優れていること、意思決定手段としての交渉(バーゲニング)に大きく依存していることなどが挙げられる。プログラム評価の支援と利用そのものに焦点を移せば、



米国政府は立法部門・行政部門のいずれにおいても多様な(支援と利用の)存在が認められる点で特徴付けられているのである。評価専門家は広範に散在し、各政府部局は内部に多様な能力を所有しており、また大学、民間会社、非営利組織および全米科学アカデミーなどの準公共機関をはじめとする多様かつ異なったタイプの評価専門家に外部委託している(Melkers & Roessner, 1997, p.58)」

連邦プログラムの評価の細分化傾向は研究開発プログラムの評価にまで広がり、このため概観して包括することが難しくなっている。

研究評価方法論についてレビューする者にとって、政治状況に起因する細分化の上に評価の公表が限られているという問題もある。この場合、評価資料の得やすさは重要な因子である。記録が公開されなければレビューは包括的になれないからである。理論上はすべての評価が入手できるはずである。連邦政府が情報の自由原則を支持しているからである。評価を公表することは非常に重要なことであるにも関わらず、実際には公表されないことが多々ある。ある機関が公表を認めるとしても、ある時点における機関のある部分に限定して取り扱う評価は委託される。こうした一般性の欠如のため評価資料は公表には適さないようになっている。公的記録も蓄積されないが、それは民間会社がたくさんの評価を行い、かつこれらの会社は普通公表に関心がないからである。インターネットがこうした状況を一変させることは確かで、本報告書に要約してある Roessner et al. や研究価値マッピングプロジェクトなどの評価は、今やインターネットのおかげで簡単に入手できる(いずれの著者もこの評価に関する論文を公表した学界関係者であるが)。

ある評価が意味を持つ期間は短いので公表記録は重要である。評価は時間と場所に依存するものなので、時間が経過し状況が変化するにつれ、古い評価への関心は消え失せ、それらは急速に視野から消えて行く。政府機関職員の異動によって評価を委託した者が去れば、機関としての記憶は失われるかもしれない。したがって政府機関職員でさえ自分たちの機関が行った評価を必ずしもすべて特定できるわけではない。真の事情を判定するための調査研究を引き受けたのは Logsdon & Rubin だけだが、もっと評価を実施してきたはずだと判っていた彼らでさえ、政府機関から一つ二つの事例しか提供してもらえなかったのである。

影響がもっと大きい評価の場合は、それほど短命ではないだろう。例えば CHI の 1990 年代後半における特許と論文間の連関(linkage)の調査研究は、特許に引用された研究の大半は公共部門から出てきたものであることを明らかにした。この研究が今だに記憶されているのはニューヨークタイムスの記事になり、かつ政府が科学予算を主張する上で使用していたためである。しかしながらこれはプロジェクト評価を厳密に述べておらず、したがってそこに含まれる意味はもっと一般性のあるものだった。評価が意思決定には影響を与えないという事実は、とくに 1990 年代初期の論評ではしばしば問題にされている。しかし評価の大半は、特定の時期に特定の情報ニーズを満たすために委託されるようである。この場合には情報がとくに好都合ならば、調査済みプログラムに関連する小サークルに限定的な評価の影響があるかもしれない。評価専門家が評価の影響を特定するのは難しい。影響は評価専門家が他の仕事に移った後の過程で現れてくるものだからである。

Kostoff の著作を除いて評価に関するレビュー論文は、1990 年代半ばに技術評価局(OTA)がなくなるまでOTAと関連のあるものばかりだった。OTA は 1985 年と 1991 年に研究評価を取り上げた2つの報告書を書き、その学問的なバックグラウンドペーパーを外部に委託した。これらの OTA 報告書が刊行された時期には、米国における評価についてのレビュー論文が学術雑誌に多数まとめて掲載された。OTA がなくなるとともに状況は変わり、Rand は 1995 年、Melkers & Roessner は 1997 年にそれぞれレビュー論文を投稿し、そして Roessner & Georghiou は 2000 年に「*Research Policy*」誌の特別号に別のレビューを掲出した。なお、これらのレビュー論文は CHI の部内記録に加えて、本報告書におけるオーバービューの基礎となっている。

レビューの対象としてはいくつかの政府機関が傑出している。基礎研究を実施する海軍研究局(ONR)は、研究評価を主題にした多くの著書を持つ Ronald Kostoff(同氏の著作はインターネットに掲載され、したがって簡単に入手できる)の本拠地である。同氏は ONR で研究評価活動を行い、それらに関していくつかの著作がある。彼の最近の大きな仕事については第 4 章に要約してある。全米科学財団も基礎研究に焦点を当て、かなりの量の評価を委託している。例えば Roessner らによる追跡調査が挙げられるが、この仕事は Abt アソシエーツのウェブサイトに掲載されている。エネルギー省でも活動しているようだが、その活動はきわめて分析されており、全体を把握するのは難しい。

近年、ATP プログラムは評価に関する動向の中では傑出しており、米国における評価風土をほとんど一方的に変えてしまった。同機関は議会の監視の下で自身の再評価を行い、そして会計検査院(GAO)からも査定を受けた。そして多くの民間の自営コンサルタントや学界に業務を委託して、その著作を公表した。この著作は主として経済学的なものだった。ATP はそのスタッフとしてエコノミストを雇用し、評価に適用する経済学的方法の開発を推進した。

研究評価領域の細分化のため研究評価専門家の完全なリストを提供するのは難しくなっている。中にはうっかり次のリストから漏れたものもあるかもしれない。

- Abt アソシエーツ—NSF のために事例研究手法を用いて評価を行った。
- フロリダ国際大学 Prof. Harvey Averch
- Barry Bozeman et al.—研究価値マッピングプロジェクトの背後にはジョージア工科大学の Bozeman やその同僚がいる。彼らは評価を社会科学理論にもとづいて大局から検証している。
- CHI リサーチ社—20 年にわたって様々な政府機関のために計量文献学による評価を行ってきた。
- Prof. Susan Cozzens—評価分野、とくに GPRA や評価に関するオブザーバー兼コメンテーターである。
- ペンシルバニア州立大学 Prof. Irwin Feller
- 会計検査院(GAO)—議会を支援し、その依頼で評価を行う。また自身でも評価業務を行う。GAO のいくつかの調査研究についての詳細な分析は Melkers & Roessner (1997)を参照のこと。
- Ronald Kostoff—海軍研究局(ONR)。計量文献学を用い、データマイニングの応用を開発している。
- 全米アカデミー—全米科学アカデミー(NAS)・医学機構(IOM)・全米工学アカデミー(NAE)・全米研究評議会(NRC)およびその委員会 COSEPUP(科学・技術および公共政策に関する委員会)。アカデミーは科学と技術に関する現在の課題について諮問する報告書を多数作成することによって国に奉仕している。これらの報告書は米国最高の科学者や科学技術者たちの委員会の仕事にもとづいている。研究評価分野

では全米アカデミーが科学コミュニティの利益を代表し、かつピアレビューの擁護と実施を担っている。

- SRI インターナショナルーこれもまた主要な評価担当機関である。David Roessner はジョージア工科大学の学者でもあり、その評価の一部を行っている。

ATP プログラムのために評価を行った主体には次のものがある。

- Drs. David Austin & Molly Maccauley **Resources for the Future**、ワシントン D.C.
- Profs. Lee Branstetter カリフォルニア大学デイビス校
- CONSAD リサーチ社
- Dr. Henry Etzkowitz ニューヨーク州立大学パーチェス校、ニューヨーク
- Maryann Feldman ジョンズ・ホプキンス大学
- Prof. Paul Gompers ハーバード大学
- Prof. Zvi Griliches (故人) ハーバード大学
- Prof. Josh Lerner ハーバード大学
- Prof. Albert Link ノースカロライナ大学(グリーンズボロ)
- William F. Long **Business Performance Research Associates**
- Haim Regev イスラエル統計局
- Research Triangle Institute**
- Mariko Sakakibara カリフォルニア大学ロサンゼルス校
- Solomon Associates、 Silber & Associates** いずれもワシントン D.C. にあるコンサルタント企業で ATP のため早くから調査を実施している。
- Manuel Trajtenberg ヘブライ大学(イスラエル)
- Prof. Nicholas Vonortas ジョージ・ワシントン大学
- Prof. Tod Watkins & Prof. Theodore Schlie リーハイ大学

## 結論

米国においてはこれまで内部評価が優勢だった。ピアレビューは、政府機関職員および全米アカデミーのどちらが行ったかに関わらず、最も広く利用され、かつ一般的に受入れられている評価方法である。外部評価は方法論的な専門知識を備えた第三者によるもので、内部評価ほど普及していないといわれ、またかなり多くの評論家の批判にさらされている。評価をとりまく状況が分析的であるために、主として学界関係者が執筆する評価についてのレビューはコンサルティング会社が行った評価の頻度を過小評価しているのであろう。過去の状況はどうであれ、現状は変わりつつある。インターネットのおかげで未刊行の著作が広範に利用できるようになり、また著作を従来よりも長期間にわたって保有しておけるようにもなって、分析化の傾向が弱まっている。外部評価を実施する機会が増え、またその勢力も衰えていないことから、この傾向は続くと考えられる。



## 第2章 ATP（先端技術プログラム）で用いられている評価方法

Rosalie Ruegg 技術影響評価研究所 (TIA)

先端技術プログラム(ATP)は、1990年の発足後まもなく、評価プログラムの開発を始めた。評価の早期開始を促した要因は、立法府の要望であった。立法府は ATP にプログラム開始の4年後に評価を提出するよう要望したのだった。別の要因は、評価が適正な管理の要であるというプログラムマネージャーの意見であった。評価の必要性に弾みをつけたのは、1994年政府業績・成果法(GPRA)の条文と外部の恒常的な情報ニーズだった。

ATP の評価プログラムは各分野の外部専門家の多大な協力を仰ぎつつ、経済評価局が開発し、先導している。学界およびコンサルティングのエコノミスト及びその他の評価者は、ワークショップ企画に参加し、また評価研究を実施するなど、重要な貢献をした。経済評価局の職員は評価プログラムの総合的な調整、研究の実施、契約の管理、進展度追跡データベースの構築、調査結果のプログラム管理者・政策立案者への提出、および評価からのフィードバックのプロジェクト/プログラム選択担当者<sup>6</sup>への提供などを行った。

ATP の評価プログラムは 10 年目を迎え、高い評価を得ている。全米科学アカデミーやその他の提携プログラムが最近個別に行った評価の結論は、「ATP の評価プログラムはその厳密さ、範囲および独立性などの点で他の米国パートナーシッププログラムを凌ぐことは明らかである」としている。評価分野における主要エコノミストである Dr. Irwin Feller は ATP 評価プログラムを「他のパートナーシップ活動のモデル」と述べている<sup>7</sup>。政府プログラムにおける浪費、不正行為および悪用などの摘発を所管する米国商務省監査長官室でさえ ATP の評価プログラムを「模範的」<sup>8</sup>と評し、賞賛した。

---

<sup>6</sup> さらに、経済評価局職員は他の ATP 職員と同様プロジェクト選定委員会やプロジェクト管理チームの一員を務めた。

<sup>7</sup> National Academies of Science, National Research Council, Board on Science, Technology, and Economic Policy, *Government-Industry Partnerships: The Advanced Technology Program; An Assessment of the Advanced Technology Program* (Washington, D.C.: National Academy Press, 2001).

<sup>8</sup> U.S. Department of Commerce, Office of Inspector General, Report on the ATP.

## 評価の折衷的アプローチ

ATP は業務に最良の方法を使う意気込みで、質と量の両面にわたるさまざまな評価方法を用いて ATP 助成金(award)の効率を査定した<sup>9</sup>。この取組み方法は最初はやや細分化されすぎるくらいがあったかも知れないが、時間が経つうちに評価計画のより大きな構造や細部の組み合わせ方が明らかになった。研究が重視する対象は様々であった(ATP のイノベーション過程への効果、イノベーターの商業的利益の向上、波及効果、共同研究の著作、評価ツールの改善、その他効果の組み合わせ)。また研究の中には対象範囲が広いものや狭いものがあった。

ATP の評価プログラムの主要目標は次の3つであった。(1) 提携プログラムの改善のために、様々なプログラムの投入、結果(output)および成果(outcome)のあいだの関係についての理解を深めること、(2)何が機能し、何が機能していないかを知るために個々のプロジェクトのパフォーマンスを追跡し、測定すること、(3)プログラムがその目標の達成に向かっているかどうかを評価するために、そのプロジェクトのポートフォリオから得られる社会的純便益の大きさを測定すること。その場合、私的便益、波及的便益、一般的な社会的便益 および ATP に帰すべき便益などに注目する。

ここでは、ATP が評価のために用いた方法の例を、その方法を用いた研究や評価プログラムの進化の過程に則して提示する。研究全体を紹介すれば読者の追加情報へのアクセスは容易になる。進行中のいくつかの研究も含め、公表済みの研究に重点を置いて述べる。

ATP は技術開発のための応用研究に資金を提供しているので、評価プログラムはより定量的な、結果指向型の経済学的評価方法を推進することができた<sup>10</sup>。ATP が用いた主な評価方法は次の5つであった。(1)プロジェクト・ポートフォリオの作成を追跡するための統計的プロファイリング、(2)資金の受領者と非受領者の調査、(3)出力データの体系的収集を伴うか、伴わないかいずれかのアネクドット研究をはじめとする事例研究、(4)費用便益、期待価値分析および感度分析な

---

<sup>9</sup> プログラムの存在全体にわたって資金供与対象プロジェクトの選定のためピアによる評価を使った。このためしばしば技術・市場評価により情報を得た。また政府プログラムの大半と同様、諮問委員会による定期的な評価や、その過程、業務、および意思決定戦略などをモニタする推進委員会も使った。それらの評価形式は本報告書の主題ではない。むしろ本報告書の主題は、資金を提供されたプロジェクトの評価である。



どの方法を使用する事例研究、(5)生産関数、回帰分析、費用指数モデルおよびマクロ経済モデルなどを含むさまざまな統計的、計量経済学的方法。また計量文献学や特許引用分析手法なども、わずかながら ATP の評価に使用されている。

### **ポートフォリオ属性の統計的特徴**

評価の最優先事項は、新たに展開するプログラムの申請者、助成金、計画、参加者および助成対象技術などの最も重要な要素を特徴付けすることであった。この情報はプロジェクト管理のために必要なだけでなく、評価にも必要だった。プログラムは適用対象となるビジネスを引き付けるのにどの程度成功したか？ 単独申請者プロジェクトに比べてプログラムは共同研究(ジョイント・ベンチャー)をどの程度引き付けていたか？ プログラムは何件の助成金を与えていたか？ ATP とそのパートナーはプロジェクトにどのくらいの資金を投じたか？ 当該プロジェクトには会社、大学およびその他の組織がどれくらい関与したか？ 関与した組織はどれか？ それらの所在地は？ 小企業は競合できたか？ どの技術分野に資金が提供されたか？ プロジェクトの継続期間は？ 外国企業には何件の助成金が交付されたか？ 何件のプロジェクトが完了したか？ 実施中のプロジェクトは何件か？ 完了前に中断したプロジェクトは何件か？ その理由は？ 以上は議会などから ATP が受けた質問の例である。

このような質問に即答するためのデータベースを構築し、維持管理することが優先課題であった。現在ではこれらのデータの収集と報告は ATP にとっては日常業務になっている。申請者、助成金採択者および技術などのデータベースは多くの評価研究にとっても情報の出発点である。

### **関係者への調査**

ATP は関係者への調査を広範にわたって実施し、申請者や助成金採択者のプログラムに対する受け止め方や実際の経験、ATP の価値に対する認識や受けた影響について情報を集めてきた。この調査方法の長所は、広範な関係者が容易に理解できるようなプログラムの効果についての定性的・定量的情報を比較的早く入手できる点である。しかし、結果のバイアスを避けるために

---

<sup>10</sup> 対照的に、主に基礎研究に資金を提供している連邦プログラムは、通常、経済・金融的方法の適用が難しく、問題があることがわかっている。

調査項目の構成には留意しなければならない。また回答率は低いかもしれない。さらに調査対象者は正直に答えるために匿名を要求するかもしれないし、またこの要求は収集されたデータの取り扱いと利用を制約するかもしれない。

ソロモン調査：プロジェクトの成果についての初期情報を得るために ATP は資金提供の1年目が終わる頃に最初の助成金採択者グループの調査を委託した。ワシントン D.C.の小さなコンサルティング会社であるソロモン・アソシエーツがこの調査を実施した<sup>11</sup>。調査は電話インタビューにより実施し、回答者にとって重要と思われる領域を特定するために、自由回答による質問をたくさん使った。回答はコード化され、表にまとめられた。調査結果は次の通りだった。(1)助成金のおかげで研究が加速された。(2)協力が重要である。(3)競争力が改善された。(4)助成金受領者が民間部門から追加資金を取付けることができたことを示す「ハロー効果」を確認できた。これらおよびその他の調査結果は短期的効果として報告された。

シルバー調査：「ソロモン調査」の数年後にもっと規模の大きい調査「シルバー調査」が実施された。実施したのはメリーランド州所在の小さなコンサルティング会社だった<sup>12</sup>。この2番目の調査は ATP の最初の3年間に資金提供を受けたすべての助成金採択者を対象にした。シルバー調査は前回調査に比べ固定式の質問が多かった。ATP のエコノミストはこのコンサルタントと密接に協力して調査を設計した。それは2部で構成されていた。1部は助成金の初期効果に重点を置き、2部は「顧客の満足」すなわち、会社はパートナーシップのさまざまな段階を通じて ATP との協力をどう捉えたかという点に重点を置いた。好ましい結果も好ましくない結果も報告された。回答者は匿名を保証され、データベースは受注者が維持管理した。調査は以下の点で ATP が参加者に及ぼすと考えられる影響度などの効果を特定した。すなわち、(1)参加者がハイリスクの研究を行うことができるか、(2)参加者が ATP の資金供与を受けなくても同じ目標、努力レベル、および速度で技術開発プロジェクトを遂行できるか、(3)助成金が呼び水となった法人研究投資、(4)協働の拡大とそれが助成金採択者に与えるプラスとマイナスの効果、(5)助成金が技術の商業化達成に与えた影響、(6)引き続き資金を受けられる魅力、(7)国際競争における立場の変化、(8)会社

---

<sup>11</sup> Solomon Associates, *Advanced Technology Program: An Assessment of Short-Term Impacts -- First Competition Participants*, NIST Contractor Report, February 1993. 同報告書は ATP 経済評価局からハードコピーでしか入手できない。

<sup>12</sup> Silber & Associates, *Survey of Advanced Technology Program 1990-1992 Awardees: Company Opinion About the ATP and its Early Effects*, NIST Contractor Report, January 1996. 同報告書はオンラインで入手できる ([www.atp.nist.gov/eao/eao\\_pubs.htm](http://www.atp.nist.gov/eao/eao_pubs.htm))。

の雇用における変化、(9) 事業の仕方の変化、(10) 非専有情報の共有が受入れられやすいかどうか、(11) その他の効果。

シルバー調査において使用された調査文書の写しはシルバー研究報告書の付属資料に掲載されている。質問票には次の 2 つのバージョンがある。(1) 単独申請者全員と、ジョイント・ベンチャー参加者(JVP)のうち事後的な商業化活動に関わる可能性が高いと特定された JVP に交付する「長文様式」、(2) 商業化事業に関与する可能性が低いと特定された JVP に交付する「短文様式」。2つのバージョンのあいだの相違は商業化と事業目標に関する質問 48 項目の部分である。

ATP の事業報告システム: ATP の評価スタッフはこれらの効果を進行に合わせて追跡する能力が必要であることがわかり、データをプロジェクト参加者からコンピュータを使って収集するシステムを部内に開発した。「事業報告システム(BRS)」と称するこの部内追跡システムのおかげで ATP はプログラムの進展を定期的に報告できるようになった。BRS は 1992 年以降に資金を供与されたプロジェクトすべてについて、ATP 資金供与期間中の進捗状況を追跡し、プロジェクト完了後も引き続きその後の展開を追跡している<sup>13</sup>。ATP は個々の会社の報告を秘密・専有情報として保管している。これらの報告には一部の会社が公表を望まない特定の事業計画データが含まれているからである。公表されたのは BRS データの集約的報告書だけだった<sup>14</sup>。

BRS はいくつかのパートから構成される。プロジェクト開始時に参加者は商業化のための技術・戦略的応用分野の計画を報告する。彼らは毎年、自分たちの商業化戦略の実施に向けた進展度やプロジェクトの短期的な経済的影響、例えば、初期販売収入、研究開発サイクルの縮小、協力効果、知的所有権の設定と保護および初期雇用創出などに関して報告する。プロジェクトの終了時には彼らは研究実績や将来の計画などを報告する。ATP 資金供与完了後の期間は、彼らは技術の商業化と普及における実績を報告する。彼らはまた技術の商業化を実施するその他すべての組織を明らかにすることを求められている。プロジェクト完了後6年以上経ってから参加者

---

<sup>13</sup> 次の文献において BRS について述べられている。Jeanne Powell in “The ATP’s Business Reporting System: A Tool for Economic Evaluation,” Paper presented at Conference on Comparative Analysis of Enterprise Data, Helsinki, Finland, 17-19, June 1996.

<sup>14</sup> BRS データにもとづく報告書の例のうち最近の報告書を見ると次のものがある。Jeanne W. Powell and Karen L. Lellock, *Development, Commercialization, and Diffusion of Enabling Technologies: Progress Report*, NISTIR, April 2000; Jeanne Powell, *Business Planning and Progress of Small Business Firms Engaged in Technology Development through the Advanced Technology Program*, NISTIR, October 1999; and Jeanne Powell, *Development, Commercialization, and Diffusion of Enabling Technologies: Progress Report for Projects Funded 1993-1995*, NISTIR, December 1997.

は3回報告する。そこでは彼らが商業上の進展をしている場合は、国に対する経済的影響をより強調するようになっている。

サイクルタイム調査：これらの対象範囲の広い調査に加えてもう少し焦点を絞った調査があった。それらは特定の問題への回答や仮説の検証のために実施されたものである。その1つは ATP 産業コンサルタントである Dr. Francis Laidlaw が、プログラム参加者の応用研究サイクルタイムに対する ATP 参加の影響を調査するために実施した<sup>15</sup>。サイクルタイムの影響は非常に興味深い。高度かつ汎用的な技術の創造と応用の促進に力を入れるという使命がプログラムに課せられているからである。

初期の調査から分かったことは、プログラム参加者の大半が ATP への参加によりサイクルタイムを縮小できると考え、かつそれはぜひ必要であると考えていることだった。しかし、実際にはその調査からは以下の点に関する詳細はわからなかった。(1)サイクルタイムの縮小がなぜ会社に重要なのか、(2)どうしたら ATP への参加によってサイクルタイムの縮小ができるのか、(3)研究開発のサイクルタイムの縮小が如何に商業化にかかる時間の縮小につながり、そして市場化がより早く可能になるのか、(4)ATP プロジェクト以外に対しても一時的な効果があるのか。Laidlaw は、1991年に ATP が資金提供した28件のプロジェクトの主要研究担当者に対し、巧みに構成した電話インタビューを実施した。回答者の大半が指摘したことは、ATP への参加の結果、研究開発面でも商業化面でもサイクルタイムが改善し、またそうした改善点は ATP 以外の技術開発プロジェクトにも応用できたという点だった。彼らは次のような特定の「ATP 的实践方法」を関連プロジェクトにも適用してみると述べた。(1)ATP プロジェクトにおいて使用または開発された方法論や過程を企業全体に応用すること、(2)過程の加速に好適な文化的バイアスの育成、(3)ATP が資金提供した技術基盤から得たあらゆる種類のアプリケーションの開発を促進すること。

---

<sup>15</sup> Francis Jean Laidlaw, *Acceleration of Technology Development by the Advanced Technology Program: The Experience of 28 Projects Funded in 1991*, NISTIR, October 23, 1997.

Feldman/Kelley 調査：調査法を用いた最近の重要な研究は、ジョンズ・ホプキンス大学の Dr. Maryann Feldman が当時 ATP エコノミスト職員の Dr. Maryellen Kelley の協力を得て主導したものである<sup>16</sup>。調査目的は、ATP 申請書の作成、ATP 申請プロジェクトへの他組織の参加(公式、非公式いずれの形でも)、選定過程の公平性に関する申請者の意見、および助成金の採択者と非採択者の競争後の経験に関する情報の収集だった。とくに興味深い点は、非採択者が各自の研究計画案を継続することができたかどうか、および採択者と非採択者はその後、他の資金供与を受ける能力に差があったかどうか、の2点だった。非採択者という統制群を使用することにより結果の信頼性は強化された(それまでの調査では採択者に対し、ATP 助成金を提供されなければ、違った行動をとったか？そしてどのように違っただろうか？というような、仮定にもとづく質問を使っていた)。

調査は1998年にATPに応募した502の申請からはじまった。この応募には合計822の組織が参加していた。1998年にATPに応募した営利企業741社に興味が集まり、そしてこの企業グループが、標本抽出枠になった。抽出標本は採択企業の100%と非採択企業の50%を単純無作為抽出して得た標本で構成した。調査文書はプリテストされ、行政管理予算局(OMB)も米国文書事務削減法の要件に照らしてそれを点検した。

インタビューはすべて6ヵ月間で完了した(1999年6月-12月)。標準的な調査手続きのあと分析が行われた。回答者は守秘義務と次のような保証を取付けた。すなわち、調査の質問項目いずれに対する回答も匿名にし、かつ特定の個人や会社の身元を明かすような形で公表しないような手段を講じるとする保証である。さらに、調査の受注者からインタビュー対象者への書簡では当該調査の実施担当組織(ジョンズ・ホプキンス大学およびボルティモア大学)が明示され、そして電話インタビューに先だって知っておくと便利な質問の抜粋が記載されていた。

### **成果(outcome)指標**

ATPは一連の成果(outcome)指標を編纂しているが、その一部は予算やGPRA報告のために使用されている。これらの指標には次のものがある。プロジェクト助成金の年間額(投入ベース)、

---

<sup>16</sup> Maryann Feldman and Maryellen Kelley, *Leveraging Research and Development: The Impact of the Advanced Technology Program*, NIST report in publication. 同報告書は2001年春にオンラインされると見られる。( [www.atp.nist.gov/eao/eao\\_pubs.htm](http://www.atp.nist.gov/eao/eao_pubs.htm) )

助成対象プロジェクトの年間件数、助成プロジェクト件数の累計、助成金の累計、タイプ別プロジェクト参加者数、完了プロジェクト件数、商業化されている技術件数、特許の出願・交付件数、出版件数など。これらの指標は主として BRS データベースから抽出されている。

### 事例研究法<sup>17</sup>

ATP は利害関係者の中に技術開発の「ストーリー」、例えば、人間的側面、研究、技術的課題、および事後の応用などに少なからぬ興味を示す者がいることを見出した。助成対象の各プロジェクトには独特なストーリーがある。事例研究法は1プロジェクトまたは1組のプロジェクトのストーリーを語る上で優れている。事例研究の中には純粋にアネクドト的なものがある。その意図するところは、プログラムとその機能の態様、科学技術、および研究が公的支援を受ける潜在的価値を理解し、判断することである。密接な関係があるのは、研究分野の進展状況を追跡する事例研究である。その他の事例研究は、プロジェクトの記載事項を統計的分析または費用便益分析などの定量的分析に結びつけるものである。事例研究の中にはプロジェクトを簡単に、スナップショット的に扱っているものもある。また深層分析を提供するものもある。要するに、事例研究法は ATP によりさまざまな方法でよく利用されてきた。ATP が委託した事例研究の6つの異なったタイプの例を以下に示す。

アネクドト型事例研究：プロジェクトのアネクドト型事例研究の多くは何年にもわたって行われてきた。ハーバード大学ビジネススクールの Prof. Paul Gompers と Prof. Josh Lerner の行った例がある。両教授はベンチャーキャピタルによる資金供与<sup>18</sup>に関するより大型の統計的研究の一環として ATP 資金提供のプロジェクトのアネクドト型事例研究を7件実施している。両教授のこれらの事例研究の目標はなぜ会社が政府資金を必要としたかを学ぶこと、および ATP 資金の獲得の前後の当該会社の状況を理解することであった。彼らはボストン地域にある新興のバイオテクノロジー会社 7 社に対し、形式にとらわれないインタビューを行った。ストーリーの主題は、新興のハイテク企業がハイリスクと見なされている技術開発を継続する資金を調達する上での困難についてである。これらのストーリーは、米国におけるベンチャーキャピタルの増加によって、ATP が

---

<sup>17</sup> 評価方法の類型学は標準化されていない。しかし方法をグループ化することは可能である。

<sup>18</sup> ベンチャーキャピタルに焦点を当てる 7 つの事例は次の報告書に記載されている。Paul Gompers and Josh Lerner, *Capital Formation and Investment in Venture Markets: Implications for the Advanced Technology Program*, NIST Contractor Report, December 1999.

資金供与の対象としてきたハイリスクだが有望な技術に関して、小規模ハイテク企業が民間資金を獲得することが容易になってきたとする考えが誤りであることを指摘している。

起源の事例研究： ATP は 1994 年から 1998 年まで、特定のプロジェクト目標を持った重点プログラムを通してその資金の大半を助成した。ATP は重点プログラムの競争的公募を 30 件編成・開催し、あらかじめ決められた問題に取り組むプロジェクトに助成金を供与した。この手法は論争を呼んだ。重点分野の選定方法に関してかなりの誤解や批判が続出した。

ニューヨーク州立大学パーチェス校の社会学教授 Dr. Henry Etzkowitz は、公共政策に資するために ATP のスタッフエコノミストである Dr. Richard Spivak と共同で、ATP の重点プログラムの一つ「医療重点プログラムのための情報インフラストラクチャー」(IIH)<sup>19</sup>の起源を追跡する事例研究を実施した。

著者たちは、会社の研究者、マネージャー、学者、団体職員および私的市民が研究のための着想(うち相当数は医療分野のもの)をたくさん提案することによってどのように ATP の「白書制作過程」に回答したかを説明している。IIH 領域における民間部門の回答がどのように ATP の期待したものと異なったかを、よく検討されたプログラムのアイデアよりむしろたくさんの断片的なアイデアを創出したことにあると説明している。しかし総合すれば、白書は潜在的には包括的な重点プログラムを描出する基盤を作り上げた。次に著者たちは IIH 重点プログラムを発展させるその後の手段として次のものを挙げている。すなわち、(1)ATP 職員による個々の論文のコレクションから包括的な白書を開発すること、(2)公開ワークショップを開催して重点プログラムの正確をもっと広範に知らせ、そしてより広範な関係者にさらなる情報提供を要請すること、(3)NIST 内部における評価、(4)IIH 重点プログラムの広報のために別途会議を開催すること、である。ATP プロジェクトマネージャーの Beettijoyce Lide とビジネス専門家の Dr. Richard Spivak は IIH 重点プログラムの追跡調査を行い、当該重点プログラムの展開中のストーリーを継続し、このため資金供与対象のプロジェクトやその将来の機会などを説明した。<sup>20</sup>

---

<sup>19</sup> H. Etzkowitz and R. Spivak, *Information Infrastructure for Healthcare: An Evaluation of a Government-Industry Technology Development Initiative*, NIST Contractor Report, October 1999.

<sup>20</sup> B. Lide and R. Spivak, *Advanced Technology Program Information Infrastructure for Healthcare Focused Program: A Brief History*, NISTIR, February 2000.

体系的出力データ集積によるアネクドット事例研究： ATP は 1997 年までには多くの深層事例研究を完了していたが、事例を体系に取り扱い、選択する際の偏向の問題をさけるような方法を探していた。採用した方法は、完了のほぼ2年後に完了プロジェクトのすべてに関するアネクドット型研究を行い、そして特定の出力データの体系的集積と全体を集約した統計分析によりアネクドット型のアプローチを補完することであった。さらに、詳細な事例研究の結果は、入手可能な場合にはプロジェクト・プレゼンテーションという形で紹介した。

完了プロジェクトの成果に関する最初の報告書は 1999 年に公表された<sup>21</sup>。それは最初のプロジェクト 38 件をカバーしている。

もっとも最近の完了プロジェクトの成果に関する報告では 50 件のプロジェクトをカバーしている。当該報告は前の 38 件のプロジェクト記事を含み、さらに新規完了プロジェクト 12 件、特許系統分析、合成実績得点および新たに集約の実績分析などを組み込んでいる<sup>22</sup>。またそれは最初および後の報告書に記載してあった完了前に中止したプロジェクトの報告に追加的な分析を加えている。

50 件のプロジェクトに対する実績格付けの基準は、新技術知識の創造における進展度、その普及における進展度、並びに技術の商業的利用の促進のためのイノベーターによる新知識の直接的な利用度など、ATP 全体の長期的成功に不可欠な「実績の3次元」といわれるものである。これら 50 件のプロジェクトのうち、実施者が弱いと格付けされた数は強いとされたものよりわずかに多く、そして大半は中程度の格付けに落ち着いている。しかし、強い実施者から得られる期待純便益だけで、50 件全体にとって確固たる実績をもたらして余りある。強力な実施者の総期待便益は事実、現在までのところ、ATP の総費用をゆうに上回る。ただし、以下の点に留意すべきである。上記プロジェクトは 2 つの異なった時期(カレンダータイム)に2つの異なったグループで評価された。プロジェクトのうち 38 件は2年前に評価され、12 件は最近評価された。技術開発やその商業

---

<sup>21</sup> William F. Long of Business Performance Research Associates (Author), (Rosalie Ruegg, ATP Project Manager), *Performance of Completed Projects*, Status Report Number 1, NIST SP950-1, March 1999. 報告書は ATP のウェブサイトで見られる ( [www.atp.nist.gov/eao/eao\\_pubs.htm](http://www.atp.nist.gov/eao/eao_pubs.htm) ) 。

<sup>22</sup> Rosalie T. Ruegg of TIA Consulting (Senior Author), (Darin Boville, ATP Project Manager), *Performance of the First 50 Completed Projects*, Status Report Number 2, NIST SP950-2, publication expected in March 2001. このあと ATP ウェブサイトに報告書の追加を掲載予定。  
( [www.atp.nist.gov/eao/eao\\_pubs.htm](http://www.atp.nist.gov/eao/eao_pubs.htm) ) 。



化には時間がかかり、そして予期せぬ進展や失敗がよく起きるものなので、これらプロジェクトの将来の更新状況によっては報告済みの調査結果が変わることがあるかもしれない。

こうした取り組みの大きな長所は、その作業によって、完了の数年後に ATP 資金供与対象のプロジェクトという大集団の動向が把握でき、また事例研究のための個別プロジェクトの選定の信頼性を低下させるような選定バイアスが避けられることにある。もう1つの長所は、最終的な報告書が政策立案者から投資家まで、多数の、かつ多様な関係者に訴えている点である。

短所は、詳細な事例研究の実施には時間や費用がかかることを前提とすると、評価対象となるプロジェクト件数の増加に伴い、完了プロジェクトの評価方法では調査の深度を制限されることである。さらに、短期間にたくさんの事例とその概要を取り上げ、集めることは難しい。したがって、早期に実施した事例はその後の事例を調査しているあいだに古くなってしまう。事例からはプロジェクトの進展状況のスナップショットが得られる。プロジェクトの最終結果は予想と違う場合もある。

協力による効率化から生じた研究費用節約額を測定する深層事例研究： ATP の最初の深層プロジェクト事例研究は、ジョイント・ベンチャー・プロジェクトに焦点を当て、5 年研究プロジェクトの開始から 2 年も経たないうちに開始された。続く追跡研究は当該プロジェクト終了時に同じ研究者が行った。

ジョイント・ベンチャー・プログラムは、当時衰退していた米国のプリント基板産業のために一連の日進月歩の技術開発を目指した。この基盤のユーザーがかねてから指摘していたのは、国内の供給者が今後 5 年以内に急速な技術進歩を遂げることができなければやがて海外の生産者に依存しなければならなくなるだろうという点だった。全国製造科学センター (NCMS) の支持と行政的リーダーシップによって、当該産業の研究能力のある会社 7 社で構成するグループおよびサンディア国立研究所は、米国生産者の競争力を高めるような研究プログラムを実施した。このプロジェクトの提案者は多くの潜在的イノベーションを特定し、それらの周辺に研究テーマを設定した。提案組織は、プリント基板の生産者かユーザーのいずれかとして直接利益を得る一方で、当該プロジェクトには米国の全産業部門の国際競争力の改善を意図したユーザーグループも関係していた。プロジェクトは 1991 年半ばから 1996 年半ばまで続いた。

これらの2つの研究はともに、ノースカロライナ大学グリーンズボロ校の Dr. Albert Link が実施したもので、研究の主題は共同研究が研究費用とイノベーションのタイミングに及ぼす影響であった。二番目の研究は最初の研究より小さい規模で行われ、新技術の採用によって会社が受ける初期段階利益を調査したが、それは研究の終了時に行われたので、イノベーションの採用はまだごく一部にかぎられていた。

Dr. Link はアンケートを使って、参加組織からその研究計画、費用見積り、競争力、およびその他の要因に関するデータを集めた。その回答は、研究開発の効率性に対して参加者たちの協力が及ぼすインパクトを推定するために用いられた。

ATP に帰すべき費用節約額の下限の推定額を確定するために、参加会社は ATP の資金供与がなかった場合でもいずれ行ったと考えられる研究活動を、まったく行わなかったと考えられる研究活動から分離するよう求められた。この操作によって当該研究の仮定にもとづく分析が行われた。ついで、彼らが ATP がなくても行っただろうと答えた一連の活動について、ATP プロジェクトが実施されなかったとの仮定の下で、実施費用を推定するよう求められた。彼らがいずれにしても行ったと見られる一連の活動について、ATP の資金供与がなかった場合の推定費用から ATP の資金供与があった場合の(会社側の)費用を差し引いた値は、ATP 助成金の最低額とみなされた。上記の推定によると研究を共同プロジェクトとして実施した結果、約 3,550 万ドルの費用節減が達成されたことになる。この推定には、上記会社が ATP 供与金がなかった場合にはまったく行わなかったであろうとした、全研究活動の 50%にあたる活動がカウントされていなかった。費用節減の概念がその場合あてはまらないからである。推定には、ATP の資金供与がなかった場合にも行ったと考えられる研究を ATP プロジェクトが加速したという、参加者の挙げた事実も含まれていなかった。

次いで研究効率の向上により新規のプロジェクト開発、プロセス開発いずれに関してもサイクルタイムが縮小できたと推定された。研究の試みは当該産業に新しい技術能力を与えた。プロジェクトはメンバー会社の生産性向上、技術の他の生産者への普及、および世界市場における米国生産者の競争力の改善などを意味すると結論された。

費用便益、期待価値および感度分析などを使用する深層事例研究： 現在までのところ、リサーチ・トライアングル・インスティテュート(RTI)の経済研究センターのエコノミストは費用便益分析<sup>23</sup>を使ってATPのために一連の広範な深層事例研究を行っている。この研究では、医療技術の査定のための事例研究評価の枠組みを開発し、それを1990年から1996年までの間にATPが資金を供与した7件の再生医療(tissue engineering)プロジェクトに適用した。この研究の対象は、がんの診断と治療、糖尿病の治療、靭帯や腱および軟骨組織などの損傷、および異種間臓器移植などの新技術だった。

各技術プロジェクトを非常に詳細に、かつ明確に点検することが必要であり、このために次の事項の分析を行わなければならない。すなわち、技術の機能の仕方、及びそれをいつどのように利用できるか、ATPは技術開発にどのような効果を与えるか、技術の開発と商業化のためにかかる過去および未来の費用、そして最初に技術を応用する場合の技術の価値、である。言い換えれば当該研究は各技術に関する研究開発・商業化・生産・医療の過程全体を特徴づけている。

便益について見ると、医療の直接費用の減少、並びに、裏付けデータが入手できる場合は患者の苦痛にかかる費用の減少の理由を、このモデルは説明している。病気の経済的負担には潜在的に次のものが含まれる。(1) 予防と治療に要する明白な支払いの形での直接医療費用、(2) 生産性損失の意味での間接的費用と無報酬の看護人が費やす資源の非明示的な価値、(3) 無形の費用、例えば、患者とその家族および友人が被った苦痛など。RTIのモデルには(1) 直接医療費および(3) 無形費用が含まれるが、(2) 間接効果は除外されている。

患者の健康の改善から得られる便益の当該研究による評価の中心には次の2つの概念がある。すなわち、「生活の質を調整した生存年」(QALY)と生活値(Value of Life)である。QALYは個人における健康便益を生活の量と質で数量化する方法である<sup>24</sup>。完全に健康な生活の1年にはQALY値1.0が与えられる。死亡はQALY値0.0が与えられる。そして不完全な健康状態の生活の1年はQALY値0.0から1.0のあいだの数値が与えられる(実際には、死亡より悪いと見なされることを示す

---

<sup>23</sup> Research Triangle Institute, *A Framework for Estimating the National Economic Benefits of ATP Funding of Medical Technologies*, April 1998. 報告書はオンライン ([www.atp.nist.gov/eao/eao\\_pubs/htm](http://www.atp.nist.gov/eao/eao_pubs/htm)) で入手できる。

<sup>24</sup> George W. Torrance and David Feeny, "Utilities and Quality-Adjusted Life Years," *International Journal of Technology Assessment in Health Care*, 5, 1989, pages 559-575.

状況の場合は負の値が付される)。特定の健康状態に対するQALY値は健康査定文献に報告されている。例えば、軽いアングナ(胸が締め付けられる症状を伴う病気の総称。狭心症など)の病気持ちの人の生活はQALY値0.90を、また重度のアングナの場合は0.50を与えられる<sup>25</sup>。こうして重度のアングナ症状での生活1年には健康状態の半年と同じ数値を与えられる。

さまざまな健康状態に付されるQALY値は、妥当な母集団に対する調査結果の平均値から導き出したものである。QALY値により専門家は健康改善を数量化できる。この場合、生活の量および質の変化を単一尺度で説明できる。

当該研究で使用する「生活値」(Value of Life)は、死亡リスクを避けるために支払うという集合的意思が示す統計的な生活の値を意味する<sup>26</sup>。完全に健康な「生活年」(Life Year)の値は統計的「生活値」から導き出せる。所与の医療技術に関連したQALYの変化は、完全に健康な1生活年の数値に適用し、患者の健康上の便益の推定ができるようになる<sup>27</sup>。

枠組みは慢性および急性両方の病気およびけがなどのモデルを次のように説明している。すなわち、(1)慢性の進行がマルコフ連鎖としてモデル化されている。マルコフ連鎖とは、患者が1つの健康状態から次の状態へ、彼らの余命にわたっての統計的差を推移することを言う。(2)急性病やけがは慢性モデルの単一期間のケースとしてモデル化される。

Bass拡散モデル<sup>28</sup>は技術拡散の過程を説明している。各技術の応用の可能性についての専門家へのインタビューにより、分析専門家は各再生医療プロジェクトの格付けと市場における最終普及状態の推定ができる。

RTIの分析専門家は、医療技術への公共投資と民間投資から得られる経済的便益を純現在価値(NPV)、すなわち、貨幣の時間的価値で調整したドルで測定する。将来の期間に得られる便益と費用は現在生じている便益や費用と比較するため割引かれる。

同研究はATPの資金供与が社会的収益を増加する次の3つの方法を確認している。(1)資金供与は研究開発を加速し、それによって新技術の導入が早くなる。早くから、またはもっと多年に

---

<sup>25</sup> 同上

<sup>26</sup> 次を参照のこと。See Josephine A. Mauskopf and Michael T. French, “Estimating the Value of Avoiding Morbidity and Mortality from Foodborne Illnesses,” *Risk Analysis* 11(4), 1991, pp. 619-631; and Michael J. Moore and W. Kip Viscusi, “The Quantity-Adjusted Value of Life,” *Economic Inquiry* 26(3), 1988, pp. 369-388.

<sup>27</sup> QALYの数値および生活データ値の利用のさらに詳しい説明は次を参照のこと。Andrew Wang, “Key Concepts in Evaluating Outcomes of ATP Funding of Medical Technologies,” *Journal of Technology Transfer*, Vol. 23(2), Special Issue Editor: Rosalie Ruegg, summer 1998, pp. 61-65.

わたって便益を得る場合は他の条件が等しければプロジェクトの NPV は増大することになる。 (2) 資金供与は研究開発の集中を強め、それによって研究開発の成功の確率を、したがって期待 NPV を高めるだろう。 (3) 資金供与は研究開発の範囲を広げ、もっと広範な潜在的アプリケーションを、したがって NPV の増加を取り込むことができるだろう。

モデルは各プロジェクトに対して次の3つの経済的成果尺度の計算を規定している。(1) 投資の社会的収益、(2) 社会的収益の1要素である投資の私的収益、および(3) 公共投資の社会的収益(つまり ATP 資金供与がある場合とない場合の社会的収益の差によって表わされる ATP 投資の収益)。各尺度は NPV および収益率を用いて計算する。分析者は感度分析を行い、プロジェクトの成果につきまとう大きな不安定性を反映させる。彼らは入力値を変えながら複数の NPV と収益率を計算する。

新しい医療技術の便益を推定するために、「ディフェンダー治療技術」、すなわち、ATP 資金供与の対象になった技術が入手できない場合に使用される最良と見られる代替治療を比較対象とする。新治療技術に関連した健康状態の悪化の具合をディフェンダー治療技術との比較で見積る場合、両者の差は患者便益の推定に利用される。例えば、網膜症、腎障害、神経障害などの3種の病気は糖尿病の一次的合併症と確認された。各病気は失明、腎臓病および末端部切断などの特徴をもつ健康状態の悪化につながる。新治療技術とディフェンダー技術を用いた結果、発病の確率が異なる場合、これは新技術の便益期待値の推定の基盤を与えてくれる。

RTI の手法はまた ATP の資金供与がない場合の状況をモデル化する反事実的仮定のシナリオを織り込み、ATP の資金供与があった場合の状況と比較する。一連の医療技術については、ATP の主たる効果は新技術の開発を加速することである。したがってとくに ATP に帰属する便益部分を推定するために、ディフェンダー技術との差は ATP の資金供与が新技術の開発を加速した期間にかぎって考慮される。反事実的仮定分析の目的は便益部分がとくに ATP の資金供与に帰属できることを確認することである。

RTI研究の主たる調査結果は、プロジェクトはそれらの私的収益よりはるかに大きい社会的収益が期待されるという事実、そしてその原因は新技術により治療を受けた患者への正のスピルオーバーが見込まれたことによるという事実だった。当該研究の結論は、ATPがこれらのプロジェクトに対

---

<sup>28</sup> 次の文献におけるモデルの説明を参照のこと。Frank M. Bass, "A New Product Growth Model for Consumer Durables," *Management Science*, 15(5), pages 215-227, 1969.

する期待収益を開発者や一般社会に拡大させる上で大きな役割を果たし、そしてそれはプロジェクトの研究開発段階を加速し、技術的成功の確率を改善したことによるとした。上記プロジェクト7件へのATP投資に対する複合的社会的収益の推定額は純現在価格で340億ドルを上回っている。

研究には大きな制約があり、その大半はモデルの範囲とその適用対象であるプロジェクトの初期状態によるものである。研究時にまだ開発の初期段階にあるプロジェクトについて、研究開発から医療成果にいたる全過程をモデル化するには大量の推定データと多くの仮定を使用しなければならなかった。実際のデータがたいへん不足しているからである。したがって、当該研究の結果は予備的なものと見なすべきである。この制約はモデル自体より評価対象のプロジェクトに関連している。モデルを事後に適用し、実際に医療適用段階に達している医療技術の評価する場合には、こうした制約は小さくなるだろう。モデル自体に関係するもう一つの制約は、ディフェンダー技術を静態的モードでモデル化することである。RTIの方法は制約があるにもかかわらず、ATPの社会福祉への期待貢献度を評価するための有益な枠組みになっている。

新しい特徴を備えた拡張事例研究モデル： ATPは費用便益分析を使ったいくつかの事例研究を委託した。これらのうちの1つは、リーハイ大学のTod WatkinsおよびTheodore Schlieの両教授によるものである。この研究は最新の手法を使って、ATPが資金を与えた一群の光通信学および光エレクトロニクス・プロジェクトに対する純社会的収益を推定することを提案している。モデル化を進歩させる提案の1つは、ATPから資金を与えられていないプロジェクトを対照群として利用することであり、その目的はATPの資金供与がなかった場合生じたであろう事態を推定することにある。この方法は仮定的シナリオとは異なる方法である。もう一つの進歩は、市場並びに顧客、競争者、サプライヤー、およびイノベーター市場の部外者などへのスピルオーバーを明示的に追跡することである。このためにねずみ算式インタビュー手法を特許・論文引用分析手法と組み合わせて使用する。

もう一つの特徴は、ディフェンダー技術が新治療技術に代替された点をダイナミックに解釈していることである。そこでは、ディフェンダー技術はATPの資金を与えられた技術と対照しながら変化するものだとしている。ただし、実際の研究が方法論的進歩をもたらしているかどうかを判断するのは早計である。

国民的便益を推計するためにマクロ経済学モデルを用いる深層事例研究： 一般に、ATPはプログラム<sup>29</sup>の成果を推定するためにマクロ経済学モデルを使用することには反対してきた。ATP予算の規模が小さいこと、米国経済の規模が巨大であることなどを考慮すると、通常は、統計的誤差からいかなる効果も分離することができない。

ATP は地域的、全国的影響を推定するマクロ経済学モデルを利用する機会がかぎられていると考えてきたが、2つの事例研究について REMI (地域経済モデリング社) モデルの利用を委託した。しかし、これら2つの事例研究を実施したのは、マクロ経済学分析の結果、プロジェクト効果をREMI 社の投入産出マトリックスに反映することができること、およびプロジェクト参加者が産業部門の比較的大きい部分を構成していることなどが確認されてからであった。

REMI モデルの基礎になっているのは、理論的、経験的關係によって経済変数を関係付ける一組の構造方程式である<sup>29</sup>。これらの關係は一般的に入手できる歴史的データから推計されたパラメータで表される。モデルはコンピュータプログラムとして実装される。モデルの利用によって地域的または全国的経済効果を推定することができる。

CONSAD リサーチ社は自動車の寸法の変動を小さくすることを目的とした自動車コンソーシアム・プロジェクトの REMI 分析を最初に行った<sup>30</sup>。この研究はまだ実施中のプロジェクトの将来便益を予測する力作だった。

NIST エコノミストの Dr. Mark Ehlen は最近 REMI モデルを使った詳細な事例研究を行った。彼は自動車部門で使われている新しい流量調節機械加工技術の経済的影響を評価した<sup>31</sup>。彼は最初に新加工技術の実行によって自動車産業の生産量と価格に生じた変化を推定した。彼は次いで REMI モデルを使って国民産出量、雇用および個人所得に対する総効果を推定した。彼は次の2つのシナリオで分析を行った。すなわち、第1は技術を採用する機会がかぎられているのではないかとする仮説、第2は採用機会はずっと広範にあるのではないかとする仮説である。

---

<sup>29</sup> REMI モデルの包括的な説明は次を参照。G. Treyz, D. Rickman, and G. Shao, "The REMI Economic-Demographic Forecasting and Simulation Model," *International Regional Science Review*, Vol. 14, No. 3, 1992, pp. 221-253.

<sup>30</sup> 「CONSAD 報告」はオンラインでは入手できないが、ATP 経済評価局で直接入手できる（電話による請求：301-975-3589 または電子メール：janet.brumbly@nist.gov）。

<sup>31</sup> Mark Ehlen, *Economic Impacts of Flow-Control Machining Technology: Early Applications in the Automobile Industry*, October 1999. 同報告書はオンラインで入手できる。  
([www.atp.nist.gov/eao/eao\\_pubs.htm](http://www.atp.nist.gov/eao/eao_pubs.htm))

## 統計的・計量経済学的方法

ATP は評価に当たってさまざまな統計的・計量経済学的方法を用いた。以下に主な例をいくつか挙げる。

イノベーションから得られる将来の消費者便益を推定するための新モデル：ともにワシントン D.C. の Resources for the Future のエコノミストである Dr. David Austin および Dr. Molly Maccauley は、別のエコノミストによる研究を利用して高度技術から得られる消費者便益を推定する新しい方法を開発した<sup>32</sup>。この新しい方法では新技術が与えてくれるサービスの質の変化を考慮に入れている。

Bresnahan<sup>33</sup>はサービスの質の改善から得られる消費者便益の推定という、これまでは問題の多い試みとしか考えられていなかった推定の実行可能性を増大した。彼は生計費指数手法を開発し、これにより革新された製品の観測価格と観測性能を、技術進歩がなかったものと仮定した場合に入手しうる最高価格や性能データと比較することを可能にした。同氏のモデルの問題点は、モデルが遡及的な評価を目指したこと、すなわち、既存のイノベーションから得られる消費者便益の推定を目指したことであった。それは ATP から資金を与えられたプロジェクトなどのまだ開発中のイノベーションから得られる消費者便益の予測にはあまり向いていなかった。

Austin と Maccauley は Bresnahan の方法を拡張して、見込み評価 (prospective assessment)、すなわち提案中の研究開発プロジェクトおよびまだ製品を市場化していないプロジェクトから得られる消費者便益の推定に応用した。二人の手法は新技術の段階的普及を認めている。モデルのパラメータは将来の、または推定のパラメータ値に不確実性を反映する確率密度関数として示される。二人はまた Bresnahan 方法を拡張して速度などの特定の製品特性に対する消費者選好を反映させ、そしてそうした選好が競争市場における当該製品の成功に影響するようにした。

---

<sup>32</sup> D. Austin and M. Maccauley, *Estimating Future Consumer Benefits from ATP-Funded Innovation: The Case of Digital Data Storage*, NIST Contractor Report GCR 00-790, April 2000. オンラインでも入手できる。( [www.atp.gov/eao/eao\\_pubs.htm](http://www.atp.gov/eao/eao_pubs.htm) )

<sup>33</sup> Tim Bresnahan, "Measuring the Spillovers from Technical Advance: Mainframe Computers in Financial Services," *American Economic Review*, vol. 76, no. 4, 1986, pp. 742-755.



図 2-1 は技術イノベーション、例えば、大量データの記憶容量の改善などから得られる消費者便益における期待利得を示している。図の左側のグラフはイノベーション前の基準を示しており、ここではディフェンダー技術だけが示される。その供給を  $S_0^{DT}$  の直線で示している。右下がりの直線  $D$  はデータ記憶の需要を示している。図 2-1 の右側のグラフは、ATP の助成によるイノベーション、言い換えれば費用削減と品質改良の組み合わせを示している。このイノベーションはその後の期に発生する。イノベーションは供給曲線の  $S_1^{ATP}$  への外側方向へのシフトとして示される。しかし、他方でディフェンダー技術も改善する。これは基準供給曲線の  $S_1^{DT}$  へのシフトで示してある。陰をつけた領域は、ある時点におけるイノベーションによる消費者余剰を示している。それは観測曲線  $S_0^{DT}$  ではなく仮定による将来の  $S_1^{DT}$  曲線について測定される。 $S_1^{ATP}$  が  $S_1^{DT}$  の右側にあるかぎり、イノベーションはディフェンダー技術に対しても改善をもたらす。

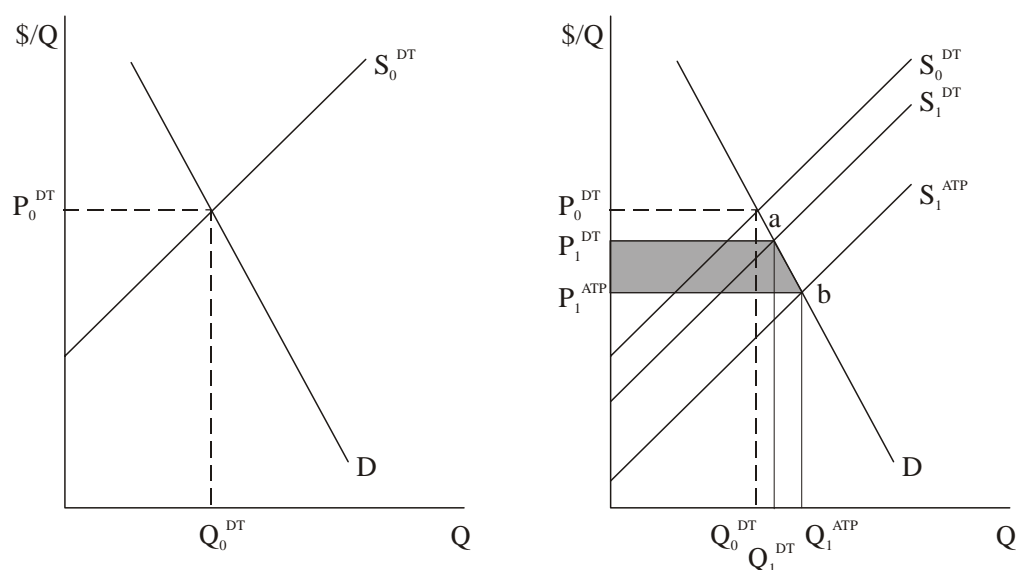


図 2-1 Derived Demand for New Technologies: Illustration of Net Surplus Change

需要曲線が計量経済学的手法を用いて推定できる場合は利得の測定は容易である。しかし、サービス部門でこれを観測するのはむずかしい。サービス部門では、実質生産量の観測は容易でないが、そういうところに限って高度技術に対して多くの需要がある。この場合、Brasnahan が先鞭をつけたトーンクビスト費用指数 (Tornqvist cost index) 手法は需要曲線の推定の必要がないため魅力的である。研究者たちは Brasnahan の手法を敷衍してこの手法は (観測できない) データの代わりに経済理論を用いるものだと説明している。図 2-1 に関して費用指数は 1 より大き

い。言い換えれば費用は基準シナリオの下ではより高く、消費者利得はイノベーションが起これば大きくなろう(総研究開発費用よりも)。

指数はイノベーションのシナリオの下での基準と比べた生活費用の変化を推定したものである。研究者たちは当該指数を構成するために既存の製品やサービスの名目単価を調整して品質・性能属性の改善に対する消費者の選好を反映させる<sup>34</sup>。彼らの仮定では、上記の影をつけた値は消費者の限界効用の逓減を反映して時間の経過にしたがって低下するとしている。

彼らはまたその手法を適用するために、ディフェンダー技術の価格を調整する。ディフェンダー技術は普通、性能が低いという属性があるため実質ユーザー費用はイノベーションの内容に比例して決まる。こうした価格調整は比較的すぐれたイノベーションの成果を達成するためには多少経費がかかってもかまわないという消費者の態度を反映している。

費用指数はラスパイレス指数(イノベーションから得られる利得を断念し、その補償を得ようとする消費者の意思を測定する指数)およびペアアシエ指数(イノベーションから利得を得るために支出する消費者の意思を測定する指数)の幾何平均として構成する。2つの指数は基準に対する相対値として測定され、理論上は2つの値は等しい。トランクビスト指数は両者のウェイトを等しくした幾何平均である<sup>35</sup>。

Austin と Maccauley は ATP 資金供与の対象であるデータ記憶技術上のイノベーションを評価する方法を用いた。第1のイノベーションは LOTS テクノロジー社が行ったものでデータ記憶容量の急激な増加を実現した光学テープ読取り／書込み技術である。第2のイノベーションはアイメーション・コーポレーションが行ったもので、磁気テープの線形走査の基本技術の開発である。この技術を使えばわずかな費用で海外の競争者がリードしている競合技術であるヘリカルスキャン方

---

<sup>34</sup>消費者価格指数(CPI)の作成のために労働統計局が用いた品質調整方法の論議については次を参照。B. R. Moulton and K. E. Moses, "Addressing the Quality Change Issue in the Consumer Price Index," *Brookings Papers on Economic Activity I*, 1997, PP. 305-366. 価格を一定期間にわたって比較する CPI と違ってここで使用される指数は単一期内で見込み価格を比較する。イノベーションを伴った将来価格とイノベーションを伴わない仮定の将来価格との比較。

<sup>35</sup> 指数の理論は、どんな指数でもそれだけでは「望ましい」特性や基準、例えば、スカラー、推移性、対称性、および比例性などすべてを満たすことはないという点を指摘している。トランクビスト費用指数は基準の多くを満たす。数学的公式については、研究者はトランクビスト費用指数を作成するためにこの説明のもとになっている次のより規模の大きい報告書を利用している。D. Austin and M. Macauley, *Estimating Future Consumer Benefits from ATP-Funded Innovation: The Case of Digital Data Storage*, NIST GCR 00-790, Gaithersburg, Maryland, 2000.

式の性能と能力に匹敵するか、それを上回ることが可能である。これら2つの技術はともに既存のテープ装置に比べて劇的な価格／性能の向上が期待できるが、まだいずれも実現していない。

研究者たちは当該方法を使って次の技術から得られる対消費者期待純便益を推定した。第1は光学テープ技術で、5年間にわたるその期待純便益は 10 億ドルを上回り、第2は線形走査技術で同期間の期待純利益は 20 億ドルであった。この分析は当該技術の早期商業化から得られる消費者便益に焦点を当て、イノベーターまたは知識スピルオーバーにより他の製造業者に生じる便益並びに第2世代製品から得られる便益は無視する。

費用指数手法は民間、公共両部門において研究開発マネージャーによる資源配分の有益なツールである。当該モデルの長所の1つは、これまで入手しにくかった情報を織り込んでいること、および同じパラメータのすべてを変えて、統一化した枠組みの中でのパラメータの値と仮定の変更の影響を知ることができるようにしていることである。新型モデルは単一プロジェクトの、一定期間の消費者便益を評価し、そしてまた民間部門や政府の研究開発資金を求めて競争しているプロジェクトの将来便益を比較するために利用することができる。

短所は不確実性や遺漏があると実際の成果が予測値と異なる場合がある点である。知識のスピルオーバーから得られる便益はモデルに含まれず、市場スピルオーバーだけが含まれる。

### **未公表の統計的/計量経済学的研究**

ATP が委託し、統計的、計量経済学的方法を使用する多くの研究が現時点でまだ終了しておらず、報告書は刊行されていない。これらのいくつかについては要約されており、ATP が評価プログラムにおいてさまざまな方法を利用していることを指示している。

- ジョージ・ワシントン大学の Prof. Nicholas Vonortas は、共同研究事業の促進のため ATP の効率性の経験的評価を行った。彼は2つのデータベース(米国法務省の国家共同研究法のデータベース、ATPのデータベース)の統計分析を行い、そして ATP が促進している共同事業が別の形態で行われうるという仮説を検証するために ATP のジョイント・ベンチャーと ATP 以外のジョイント・ベンチャーを比較した。彼は、その仮説は成立せず、ATP はジョイント・ベンチャーの成立を促進していると結論づけた(未公表 ATP 報告)。

- UC デービスの Lee Branstetter および UCLA の Mariko Sakakibara の両教授は、計量経済学的方法を使ってプロジェクト参加者の研究開発に関する共同研究事業の影響を評価した。同研究は研究開発の成果の尺度として与えられた特許を使用した。彼らは日本の特許データを使ってモデルを検証した。次いでわずかな ATP 特許データを使って ATP から資金を供与された共同研究事業が企業特許に与える影響を調査し、プラスの効果を発見した(未公表 ATP 報告書)。
- ハーバード大学の Prof. Zvi Griliches は、最近亡くなったが、当時はイスラエル、ヘブライ大学の Mauel Trajtenberg およびイスラエル統計局 Haim Regev 両氏と ATP のために計量経済学的研究に協力していた。彼らは政府研究開発補助金が企業業績に与える効果の分析方法を開発した。そのために ATP プログラムより歴史がはるかに長く、データも多いイスラエルの類似プログラムの実験データを使った。同研究は入手可能な長期的なデータを使って政府研究開発資金の企業への供与が企業の成果と生産性に与える影響の分析がうまくできることを例証した。研究の目的は、政府補助の研究開発と、企業の私的資金調達による研究開発とでは、企業業績に与える影響が異なるかどうかを検証することであった。第2の目的は、さまざまなプログラムのメカニズムや手段がどのように機能するか、およびそれらがプログラムの目標にどのように貢献するかを分析することだった。さらに広範な目標は、それまでの政府資金供与の生産性の研究方法に改良を加えることであった。それまでの研究は政府資金供与の種類を区別しない傾向があったからである(未発表 ATP 報告書)。

## 要約と結論

Prof. Irwin Feller は全米科学アカデミー後援の最近のワークショップでトルストイを引用してこう述べている。「うまくいっても人は満足しないだろうが、失敗すると不機嫌になるのは確実だ」と。彼はこれに評価者バージョンを続けた。すなわち、「悪い結果を正確に評価すれば確かにプログラムを終わりにするだろうが、よい結果を正確に評価してもプログラムを救えるとは限らない」<sup>36</sup>。したがって、全米科学アカデミーが、ATP は「非常に厄介な目標を追求する連邦プログラム」であり、

---

<sup>36</sup>National Academies of Science, National Research Council, Board on Science, Technology, and Economic Policy, *Government-Industry Partnerships: The Advanced Technology Program; An Assessment of the Advanced Technology Program* (Washington, D.C.: National Academy Press, 2001).

ATP の評価作業機関が他のいかなる研究資金供与機関よりも立派に報告書を作成することができると結論づけたとしても、結局は ATP に対する政治的反対勢力が顔を利かすことになるかもしれない。それでもやはり、過去 10 年間に ATP のために評価研究を行った機関は、評価に適用する経済学的方法における最新技術を向上させ、その後の評価研究のために高い水準を提供した。

ATP がとくに強力な評価プログラムを開発できた理由はいくつか考えられる。その1つは ATP の評価プログラムは他の政府機関の多くとは異なり、GPRA に反発せず、むしろ正面からこれを受けて立っている。ATP はそのマネージメントがそうした方が有益であると考え、その評価プログラムを立てた。その時期は GPRA や ATP に対する批判よりも前のことで、1994 年にはじまっている。評価プログラムの主たる動機は、プログラムの構築・運用担当者がさまざまなプログラムの投入、産出および成果の間関係をできるだけ理解してプログラムをもっと良く機能させたいと考えたことであった。同時にプログラム批判者が ATP のすべての評価結果をきびしく精査するはずだと考えたため、ATP はプログラムをきびしい目でみる姿勢を強めた。

ATP 評価プログラムの長所の第2の理由は、その担当部門が評価のベストプラクティスに従う点にしっかりと注意を払っていることだった。これは一つには仕事にベストを尽くすことを重視する NIST の風土から影響を受けていた。ベストプラクティスに従うことの1つの側面は、ATP に対する信頼性を維持することであった。技術評価分野の主要エコノミストなどから関心を勝ち取り、そして高度な最新水準の評価研究に資金を供与しようという意図を示すことによって、ATP は重要な信頼できる評価プログラムの構築に必要な人材を取り込むことができた。この努力には評価部門に「広報」活動を持ち込もうとする他の担当者による忍耐強い努力が含まれている。それはまた提案され、完了した評価研究のピアレビューを確保する作業が含まれている。

評価者は ATP の評価にいくつかの手法、とくに費用便益分析や統計的分析を使う調査方法や事例研究法を徹底的に利用した。こうした努力の結果、事例研究法の豊富な記述が多様な関係者のあいだに技術イノベーションへの関心を呼び起こし、そして事例を統計的分析、費用便益分析などの定量的ツールに結び付け、成果の有益な証拠をもたらした。それは評価研究のために反事実的仮定や対象群を標準慣行として利用した。ATP が成功するために満たすべき特別な検

査を設けることによって、評価プログラムは内部に成功基準意識を醸成した。そして理論的に正しい目標に対して実際に成功したのは何時だったかを知るための基盤を提供した。

最後に、ATP は評価の肥沃な土壌を提供した。評価者に協力する内部エコノミストや学者およびコンサルティングエコノミストは、ATP が最新のイノベーション知識を前進させ、官民のパートナーシップ・プログラムを通して経済的便益のために促進する方法をよりよく理解するための実験室を提供する可能性が十分あることを理解した。

## 第3章 計量文献学

Diana Hicks & Francis Narin CHIリサーチ社

### 序文

計量文献学データからは、ファンディングや学生数など他の研究開発に関する定量的データとは異なる、特有の付加価値が得られる。定量的データが揃っていれば総計を出すことができるが、計量文献学データからはさらに引用が得られ、かつ引用には様々な用途がある。評価のためには、その対象になっている研究の引用される総数がとても重要である。引用と言えば連関(linkage)を思いつくが、評価に有用な連関のデータには、論文と論文、特許と特許、および特許と論文、という3つの種類がある。連関の情報は計量文献学データに特有のものであり、引用の取り扱いに関する方法論的事項は広い層から認められているわけではないので、理解不足から批判が起こる場合もある。本章は評価に関連する3種の引用データすべての背景や用途を述べ、方法の発展について説明し、妥当性確認の研究に簡潔に触れて、長所と短所および評価に用いる際のベストプラクティスの要素を検討する。

### 歴史と妥当性の確認

#### 科学引用分析の起源

引用を使った方法で科学者の研究のインパクトを測るというアイデアの起源は、次に引用するGarfieldの論文(*Science*, 1955)であろう。

「1873年以來、シェパード・サイテーション社(コロラド州コロラド・スプリングス)は非常に貴重な研究ツール「*Shepard's Citations*」を法律専門家に提供してきた。少し前から私は、科学のための引用コード開発に関与するようになった。科学索引を効率よく機械入力するのに必要だったのだ。明らかにこれは、特定の研究の重要性とそれが当時の文献や思想に与えた影響の

大きさを評価する歴史的な研究にきわめて有益である。「インパクト・ファクター」というものは、科学者の出版物の件数を漏れなく数え上げることよりも、はるかに多くの示唆を与えてくれるだろう。」(Garfield, 1955, p.108-109)

1960年代初めに Garfield が作り上げた科学引用索引 (*Science Citation Index*) はその後、年間に科学雑誌 5,000 部以上、論文 50 万件以上、引用 500 万件以上を取り上げるまでになっている。

同氏と共同研究者らは引用データを利用して個々の論文のインパクトを測定できることには十分気付いていたが、科学引用データが評価に広く取り入れられるようになったのは全米科学財団 (NSF) が 1972 年に「*Science Indicators*」を初めて作成したときであった。Narin と CHI リサーチ社 (当時はコンピュータ・ホライズン社と称していた) の共同研究者は *Science Citation Index* データを利用して、国内にも国際的にも通用する科学のパフォーマンス指標を創り上げた。彼らは出版物の集計、そして最も重要である当該出版物の引用頻度の集計を利用し、国内科学のパフォーマンスに対する初の計量文献学的指標を作成して上記報告書に用いた。そうした作業は隔年で継続した。ある年には新しい標準指標を作成し、2 年後にまた新指標を開発するといった具合だった。

1970 年代後半から 1980 年代前半にかけて CHI は研究評価のための引用分析手法を開発した。この開発の多くは米国立衛生研究所 (NIH) のためになされた調査の一環であり、第 4 章に概説してある。この仕事は当時、評価のレビュー論文などに広範に報告された。

科学引用分析のために開発された手法を特許引用分析にも使えるように、引き続いてさらなる開発が進められた。こうした手法や知識を科学分野から特許分野に移転する動きを加速したのは、1980 年代初期の米国特許局の特許書誌データ (bibliographic tapes) の公開だった。このテープには発行済みの米国特許の標題や抄訳だけでなく、それらの特許のとびら頁にある「引用された文献」のすべてが採録されている。CHI はこのデータベースを使って分析的特許引用指標データベースを構築した。このデータベースに使用した分析手法はその後 20 年にわたって大幅に改善され、より正確さを増している。当初の特許引用指標は、初期の *Science Indicators* と同様に国内的パフォーマンスおよび国際的パフォーマンスの尺度として考えられていた。後にこのような政策レベルからはるかに厳密な国家戦略レベルに高められ、企業における技術の生産性、イ



ンパクト、重要性などの尺度を扱うようになった。さらには個別技術を扱うまでに精巧なものとなって、個々の技術の開発を追跡したり技術分析したりするのに使用された。この仕事は特許化技術の産出に興味をもつ政府機関の利用するところとなった。

引用分析の最近の開発は特許と論文間、または連関(linkage、すなわち科学技術連関)分析の分野で進んでいる。CHI は 1980 年代初めに、特許における引用の対象が特許から科学的論文を含む非特許資料へと拡大した事実に気づき、特許間の引用分析手法をこの新分野に拡張しようとした。それ以来、特許による論文の引用量が劇的に増大し、研究成果の新指標の利用ができるようになった。言い換えれば特許における論文の引用頻度が増大したのだった。1990 年代後半に広範に報告された CHI の分析によると、特許が引用した論文の 75%は公共部門の機関が作成したものであることがわかった。CHI は政府機関のためにも多くの連関(linkage)研究を行ったが、これらは第4章に要約してある。

論文同士の引用から特許同士の引用そして最後に特許と論文間の引用にいたる計量文献学の発展には数十年がかかった。この間、方法論的イノベーションには引用頻度の大きさと指標に付与できる意味とのあいだの相関関係を検証する研究が伴った。これらの研究は次節で見て行くことにする。

### 科学引用分析の妥当性の確認

科学引用分析の妥当性を検証する研究の大半は、相関分析の手法を用いて行われた。これは引用頻度の大きさと、その他多くの入手はできるが未だインフォーマルな科学業績・質の尺度とのあいだの相関関係を研究したものである。無論、科学引用分析を扱った論文は何千とあり、そのうち数百は上記の妥当性確認の概念を直接扱っている。早期の論文のうち約 24 件は、CHI の専門書『評価のための計量文献学』の第 V 章「非文献尺度との相関関係」で分類し、確認と議論がなされている。なお、この著作は研究評価における計量文献学の初期の利用状況を全米科学財団のために報告したものである(Narin, 1976)。

初期の妥当性確認手法は、研究評価に用いられているすべての研究を扱っていた。この手法は国家、機関、研究グループおよび個人の成果を対象として出版物と引用尺度とのあいだの相

関関係を取り上げていた。例えば、国家レベルでは、Derek Price の非常に初期のある著作は、国家がほぼ国内総生産(GDP)に比例して論文を出版していること、つまり人口や土地面積などではなく、その経済規模に比例して論文を出版していると述べた。CHI はずっと後になって、こうした傾向は技術にも持ち込まれるようになり、外国の投資家は一般に GDP で測定した彼らの国の経済規模に比例して米国特許システムで特許を得ていると述べている(Narin, 1991)。

機関レベルでは、引用手法は広く大学の学部の格付けに適用された。米国における格付けは一連の報告書の中で体系的に行われた。その報告書には比較的多数の年配の大学人による主要学部の相対的な格付けもなされていた。1978 年に出版され、1980 年に再版になった論文の中で CHI はこう述べた。「これらのピアレビューによる大学の格付けは論文による格付けとよい相関関係にあるだけでなく、引用データを利用すればこうした相関は大幅に強まる」。言い換えれば、ある大学の論文数とその引用頻度の組み合わせにもとづいた格付けは、論文頻度だけにもとづいた場合よりピアレビューによる格付けとの相関関係が強い(Anderson, Narin and McAllister, 1978)。

引用手法はまた個人の研究業績の重要度の測定にも用いられる。これらの手法は、科学コミュニティ内の著名度や評判などその他の尺度とともに、大学学部の引用データと学科の地位や高名な賞との間の相関関係の研究に使用された。このようなケースの多くから、科学における生産性や品質の独立尺度と論文から開発した尺度、とくに引用手法とのあいだには、一般に 0.5 から 0.7 の範囲の強い相関関係があることが示されている。

頻度の高い引用の重要性を最も魅力的かつ顕著に示しているのは、科学分野のノーベル賞受賞者の計量文献学的特性に関する一連の論文である。研究の質とノーベル賞を論ずる論文で Inhaber はこう指摘している。「引用頻度で測定したノーベル物理学賞受賞者の研究の質は他の科学者の場合より 10 倍も高い」(Inhaber & Prednowek, 1976, p.34)。

## 特許引用分析の基礎

米国特許が登録される場合は大抵、そのとびら頁に7つか8つの「米国特許に引用された参考文献」が掲載されている。これらの引用文献は、発行された特許を引用された(先行する)技術に関連づけ、登録されたばかりの特許の請求範囲を制限している。それらは重要な関連技術がす

で存在していることを指摘し、そしてそれ故に引用した特許の請求範囲を制約している(または請求範囲に対する反論として引用されている)。

特許の引用はある重要な点で科学論文における引用と異なっている。すなわち特許とびら頁の参照文献は特許審査員が記したものか、出願者やその弁理士が提案して検証対象の技術の専門家である特許審査員が承認したもののいずれかである。こうした経緯により、特許における引用が論文の場合よりも強力であることは間違いなく、かつ当該特許の分野との関係もより深いと考えられる。

この参照パターンを逆転させ、所与の特許のその後の引用をすべて一覧表にすると、特許引用分析に使用される基本的情報、すなわち、所与の特許がその後の特許に何回引用されたかについての情報が得られる。これらの頻度分布は非常に偏りがちである。すなわち、ほんの数回しか引用されない特許が多く、何度も引用される特許はごくわずかしかない。特許の半数は5回以下しか引用されず、わずか1%が48回も引用されている。

#### **特許引用分析に関する妥当性確認の研究**

本節では様々な調査研究を取り上げるが、それらは特許の平均引用頻度ではなく特許の重要性を表す他の指標、例えば、同僚科学者の意見、技術の優秀さに対する授賞、および米国裁判所が与える「先駆的な特許」という法的地位などとの関連が強い。特許引用を重要な特許を探す手段として考えた最初の論文は非常に早期に IBM の Reisner によって書かれたものである。それは、重要な特許を探すために引用分析法を利用する試みであった (Reisner, 1965)。彼女は参照項目を1つの特許から別の特許へと追跡することによって、彼女が探し求めている 60 の重要な特許のうち 47 を見つけた。

登録済みの米国特許すべてを扱う完全な引用データは、1975年に初めて得られた。その翌年に特許・商標局は、第6回技術評価予測報告書において引用頻度が最も高い特許を表にまとめ、「特許文書が引用される頻度はその技術的重要性の尺度になるだろう」との見解を示している (PTO, 6<sup>th</sup> Report, 1976)。1970年代半ばに英国において Ellis, Hepburn and Openheim の3人が特許引用ネットワークで実験をし、数組の特許から参考文献を追跡して、一群の技術の中から鍵となる発見と転換点を特定できるかどうかを調べようとした (Ellis et al., 1978)。この分析は比較的成功を収めた。非常に重要な出発点になった事例について、鍵となる発見がいくつかの特許

群の内の一つに特定できたからである。他方、技術が比較的長期にわたって開発された分野では技術は鍵となる特許群を持たず分散する傾向となった。

最初の比較的本格的な特許引用分析の研究は、全米科学財団の後援の下で CHI リサーチ社が実施した (Carpenter, 1981)。その研究が 1970 年代後半に提案された時、全米科学財団の科学指標部門は、特許引用にもとづく技術指標を当時の『*Science Indicators*』の報告書で使用されていた科学論文指標のグループに加えるべきかどうかを検討していた。NSF は CHI に委託して、重要な発見に関連した特許が平均的な特許よりもっと頻繁に引用されているかどうかを知るための調査を行わせた。『*Industrial Research and Development*』誌が設けた IR-100 賞を受けた製品の基盤になっている主要特許を決定する試みにより、一組の重要特許が得られた。この賞は、

「年間に開発された 100 件の重要な新技術製品とその発明者に与えられる。『*Industrial Research and Development*』誌の著名な編集諮問委員会は何千という候補の中から最も重要で、ユニークで有益な製品を 100 点選択する。広範な地方および国内のプレスやテレビ局が授賞候補と授賞式を取材したため、IR-100 の受賞は応用研究・開発分野で最も切望される業績となった。( *Industrial Research & Development*, 13, p.3, December 1980) 」。

調査には特許が引用されるまでのしかるべき時間を保証するために(最近の受賞対象を避けて)1969 年と 1970 年の授賞に関連した特許が使われた。重要な特許 100 件のセットと、比較するための対照特許 102 件のセットが選択された。調査からわかったことは、IR-100 特許が平均で対照特許の2倍引用されていること(IR-100 特許の引用頻度 4.9 に対して対照特許は 2.0)、および IR-100 特許の中で 10 回以上引用されているものが 17 件もあるのに対照特許の中で 10 回以上引用された特許は4件しかなかったことであった。IR-100 特許の引用頻度が対照特許よりはるかに多く、かつ非常に多い可能性が高いことは明らかである。この差は IR-100 のセットに引用頻度が多い特許が含まれていることによる。

この調査のあと、『*Science Indicators*』(後に *Science and Engineering Indicators* と名付けられた)報告書に特許引用指標が加えられ、それ以来科学指標の利用は増加した。やや異なる方法によって同じ結果を得たのが、Trajtenberg による『引用をつけて何になる("A Penny for Your

Quotes”』(Trajtenberg, 1990)という一風変わった題の研究である。Trajtenberg は CT スキャナー(コンピュータ断層撮影装置)の進歩に関連した特許引用パターンを分析し、引用にもとづいた特許指標と、それとは独立の CT スキャナーに関するイノベーションの社会的価値という尺度とのあいだの密接な関連を指摘した。とりわけ重要なのは、「重要度は引用回数に対して非線形(増加)であり、また引用の情内容的内容は限界的に上昇する」という優れた結論が出たことである(p.172)。この結論は、引用頻度の高い特許はとくに技術的重要性が大きいという考え方を直接裏付けている。

CHI リサーチ社はイーストマン・コダック研究所と協力して同研究所の技術および競合する他社技術の分析に特許引用データが使用可能かどうかを調査した。コダックはそれとは別に、産業界の研究所内において知識を持つ同僚が特許の重要性を評価する際に、頻回な特許引用を評価と関連付けるかどうかの確認を望んだ。調査は注意深く企画され、コダックの中核を担うハロゲン化銀技術に関する 16 件の特許を研究所上級スタッフ 20 人に依頼し、評価を求めた。どの特許も少なくとも3~4の異なった組合せの中に入るように組合せを重複させ、特許の格付けを相互に表にまとめることができるようにした。コダックの評価者は上級知的所有権スタッフ、研究所上級管理者および研究所上級科学者などであった。科学者が自らの特許を格付けしないように、割当ての特許を選択した。各人は与えられた特許について、最も技術的重要性の高い特許から最も低い特許までの格付けを求められた。結果は、ある特許の引用が1回であれ、2回であれ、3回であれ、引用頻度は同僚の格付けにおいてそれほど重要ではなかった。しかしながら、5回以上引用されている特許、すなわち、引用頻度が比較的高い特許はコダックスタッフによってとても高く格付けられた。同調査における 15 人の回答者のうち8人は引用頻度が最も高い特許に最高の格付けを与えた。2項分布モデルによれば、この確率は 0.0002 である。

引用頻度の高い特許の重要性に関する最近の裏付けは特許局の内部から出された。、引用頻度と特許局の認可とのあいだに強い関係があること、先駆的な特許の引用頻度は非常に高いことを示した。CHI は次の3つの異なる特許区分の引用頻度を調べた。「全米発明家の名誉の殿堂」に記載された特許、商務省が米国建国 200 周年祭のために作成したリストに記載された「歴史的意義(Historical Significance)」のある特許、そして連邦地方裁判所が先駆的と判定した特許などである。

『全米発明家の名誉の殿堂』は米国がその特許制度を通じて育成する偉大な技術進歩を考え付いた個人に捧げられる。

発明家は『全米発明家の名誉の殿堂』財団の選定委員会により殿堂入りの選定を受ける。選定委員会は全米の科学・技術団体の代表者で構成される。委員会のメンバーは毎年、候補者の中から最も優秀な発明家を選ぶ投票を行っている。

選定委員会は投票に際して、次の3点を考慮する。すなわち、候補者の発明が米国特許の対象になっているかどうか、当該特許の国家の福祉への貢献度、および当該特許が科学と有用な技術の進歩を促進する度合い。」(全米発明家名誉の殿堂、1993)

調査により、殿堂入り、歴史的意義、先駆性(のいずれか)に当てはまる特許の大半は引用頻度が平均よりかなり高く、歴史的意義だけを認められた特許は引用頻度の平均を下回っていることが分かった。先駆的特許は特定年の平均的特許より5回も多く引用されていた。

#### 引用頻度が高い特許における資金的・経済的相関

特許引用率と技術の質との連関(linkage)を示すもう一組の例は、経済的・資金的データと特許引用回数とのあいだに明らかにされている関係から推察される。

1980年代初めまでは大規模な特許データを入手することができたので、ハーバード大学のGrilichesと彼の同僚、および国立経済研究局(NBER、米国)は特許の経済的重要性を調べる一連の定量的研究を開始した。Grilichesは1981年の論文で、会社の市場価値と、過去の研究開発支出および特許件数によって形成される「無形」資本とのあいだに、顕著な関係があることを発見した(Griliches, 1981)。

Narinと彼の同僚は1987年に米国の製薬会社18社のグループを調査し、会社が取得した特許の数、特に引用頻度の高い特許を有しているかどうか、それら会社に対する同業者の見解や製薬会社の販売額および利益と相関関係があることを示した(Narin, Noma and Perry, 1987)。その研究から、(1)引用頻度の高い特許は、スミス・クライン社の製剤タガメットのように、経済的に重要な発明をめぐって生まれる傾向があること、(2)これらの重要な技術的事象は、当該産業において会社の販売額と利益の増加につながっていること、が非常に明瞭にわかった。

1992年と1993年に『*Business Week*』誌は、10業種にわたって大企業を格付けするCHIリサーチ社のデータを使って、2つの特許スコアボードを公表した(Buderi et al., 1992; Coy & Carey, 1993)。これらの特許スコアボードは、このような着想が産業界に直接導入された最初の2つである。これによって分析専門家が会社の業績と技術力との間の関係を見ることができるようになった。

NBERに関係するエコノミストは今や多様な方法で特許引用手法を利用しており、会社間または大学と会社間から生ずる研究の波及効果の調査や、成功会社の特性の研究を行って、一般に特許引用は統計的な意味でインパクトの高い技術と等しいという概念が受入れられたことを論証している。Jaffe, Trajtenberg and Henderson(1993)の論文はこれらの文献と関連している。

最近のNBERの報告書『市場価値と特許引用: 初期的観察』は「引用で加重された特許ストックは、特許ストックそのものよりも市場価値と高い相関関係があり、これは引用頻度の非常に高い特許を保有している企業に高い価値が置かれていることが主要因である」という事実を発見している(Hall, Jaffe, and Trajtenberg, 1998)。

ハーバード大学のF.M.Schererおよび欧州とCHIの同僚による最近の報告書では、米国とドイツの特許を得た発明(この発明に関する収益性情報、すなわち特許の私的価値はすでに明らかになっている)の例を取り上げている(Harhoff, Narin, Scherer & Vogel, 1999)。彼らは当該特許をその有効期間である18年間、ドイツで登録しておくために必要なすべての費用を支払っている特許だけを取り上げ、次いで当該特許の所有者に特許の資産価値に関して質問した。要点は「この特許を独立した第三者に1980年に売却するとしたら、最低いくらだったら売却するつもりか?」であった。ドイツの特許制度では、最高の価格区分に入る2つの特許はその他の特許よりはるかに引用頻度が高かった。米国特許制度では、推定価値が2,000万ドル以上の特許の引用頻度は、それより低い推定価値の特許よりはるかに高かった。

Deng, Lev and Narinは、特許引用指標と、研究開発予算や株式市場実績を含むさまざまな金融指標とのあいだの関係を調べた(Deng, Lev and Narin, 1999)。とくに彼らは、所有している特許が(一定期間)平均引用以上の指標やサイエンス・リンケージを有する会社がかかなり高い時価評価率や株式収益を、いずれも現在と今後長年にわたって有する傾向があることを発見した。

CHIが特許指標を使って実績の高い株式ポートフォリオを選択する方法を開発したことは驚くことではない。CHIは特許引用手法を利用して、株価がその後数年間に上昇するとみられる会社を特定する方法の特許を得た。

## 特許・論文間の計量文献学

CHI リサーチ社は 20 年以上にわたって米国特許による科学論文の引用を分析し、科学と技術とのあいだの連関(linkage)指標としてそれらの利用可能性を調査してきた(Carpenter et al., 1980)。特許における科学論文引用に関する最初の調査研究では、科学論文のために開発した計量文献学的分析の拡張が求められていた。そこでは科学研究と特許取得が密接に関連する 2 つの急成長分野、プロスタグランジン(ホルモン物質)とガスレーザーが調査された。当該調査研究の目的は、技術開発と研究とのあいだに密接な関連があるとみられている分野の特許が、科学論文を多く引用するという密接な関連を示すかどうかを発見することであった。その結果、これらの科学関連分野の特許が実際に相当量の科学論文を引用していることがわかった(Carpenter & Narin, 1978)<sup>37</sup>。その後の研究において 19 人の優秀な研究開発マネージャーに 24 の技術を科学との依存関係で格付けするよう求めた。それらの格付けは論文の引用率とよく一致していた(Carpenter & Narin, 1983)。CHI はこの有望な手法を追究し、事例研究から体系的なレベルの分析に移行するのに必要なデータベースを構築した。この結果、今日、特許による論文の引用は 1978 年当時より強力で広範な基盤を持つにいたった。最も科学と密接な分野は生物工学で、この分野では特許が平均 20 点の科学論文を引用している<sup>38</sup>。一方、とくに研究集約的ではないと見られている分野、例えば、産業機械、工具、または繊維およびアパレルなどでは、特許はそれほどたくさんの科学論文を引用していない<sup>39</sup>。CHI は、この種の引用データは科学技術連関の定量的比較分析を可能にする上で非常に優れていると論じている(Narin & Olivastro, 1992<sup>40</sup>)。

1997 年には Narin と同僚が、米国特許による米国論文の引用に関して影響力の大きい研究を公表した。その報告書は 1987 年-88 年と 1993 年-94 年に発行された 39 万 7,660 件の米国特許のとびら頁に掲載された 43 万 226 件の非特許参照(NPR: non-patent references)を分析した。43 万件の NPR のうち、約 24 万 2,000 件は科学的参照、すなわち科学雑誌論文、学会およびその他の科学出版物の引用と判断された。これら 24 万 2,000 件のうち約 17 万 5,000 件は科学引

---

<sup>37</sup> 審査官および申請者言及事項のいずれもが検証された。プロスタグランジン特許では、SCI ジャーナル記事は 1 特許当たり 8.1 の率で言及された。8.1 の内訳は 0.5 (審査官) + 申請者 (7.6) である。ガスレーザーでは、この率は 1 特許当たり 1.9 で内訳は 1.1 (審査官) + 0.8 (申請者) である。

<sup>38</sup> 1999 年に会社に発行された米国で発明された特許。審査官(とびら頁)による引用のみ。

<sup>39</sup> 1999 年に会社に発行された米国発明特許に関する科学文献の審査官(とびら)による引用：産業機械と工具 0.22、および繊維とアパレル 0.37。

<sup>40</sup> この報告書にはその時までの CHI の連関(linkage)分析の詳細な検討が含まれている。



用索引 (*Science Citation Index*) に記載された論文の引用であった。SCI に採録されている論文と合致する参照はさらに、引用した特許の発行年に先立つ 11 年間に発表され、かつ米国人著者が記載してある論文だけを含むように絞られた。4 万 5,000 件の論文が残った。これらについて、著者が謝辞に記した研究資金の提供者をすべて図書館で調べた。

同調査研究は、科学に関連のあるすべての産業領域にわたり、会社の大小によらず、公共的な科学が米国産業を支援する上できわめて重要な役割を果たしており、かつ米国の技術進歩の基本的な柱になっていると結論づけた。さらにデータからは、ハイテクに貢献する科学は主流であることがわかった。つまり、それはとても基本的かつとても新しく、影響力の大きいジャーナルに公表され、主要な大学や研究所の著作であり、そして NSF、NIH および国防省、エネルギー省などの省並びにその他の公共機関、財団の支援を受けている、ということがわかった。

当該研究はニューヨークタイムスで報道された。関係者によると、それはホワイトハウスや議会での予算審議に大きな役割を果たした。最近ではオーストラリア政府のために、科学や技術における政府の意思決定に影響力をもつ調査研究が企画されている。この研究は第4章に説明してある。

## **長所と短所**

### **論文間の計量文献学**

計量文献学的分析が複雑な点はたくさんあり、しかも多様である。問題点の中には技術的要因によるものがあり、これらは注意深いデータ編集により解決できる。また科学者の引用慣行に起因するものもあるが、これらは適切な研究設計や解釈を行えば対処できる。

科学論文を測定しようという試みは現在、米国フィラデルフィア州の科学情報研究所 (ISI) のおかげで可能になっている。この研究所は多数のジャーナルにおける論文の詳細情報を編纂し、それらを『*Science Citation Index*』(SCI) として利用できるようにしている。SCI を利用する際の技術的問題は、取り扱い範囲、一貫性および筆頭著者の索引など 3 種に分けられる。ISI は非常に多くのジャーナルを精査するが、その多くは対象外とされ、書籍やその参考論文リストはほとんどデータベースから除外される。ISI は多数のコア・ジャーナルを採用しているが、それらは世界中の主要論文の大半を占めているといわれる。この結果英語以外の論文に対して偏りが生じている。

つまり、日本人の目から見れば、SCI の引用や出版物の集計は国際指向の論文に限って公平に評価している。

一貫性の問題はデータの誤記録に関するものである。誤りをおかすのは、論文の著者自身、ジャーナル、またはISIであり、その結果引用や論文がデータベースへ不正確に入力されることになる。こうした問題に十分な注意を払えば、例えば、引用と被引用論文を照合するのに高度なアルゴリズムを踏めば、計量文献学的分析は非常に正確になりうる。

第3の問題は筆頭著者だけを引用することから生ずる。このため、「D Smith」を探しても Smith の論文をすべて探し出すことにはならない。また、名前の頭文字と姓を同じくする別人の著作の引用を含むことになろう。最も満足すべき解決策は完全な参考文献リストから作業し、そして引用を完全な参照論文と照合することである。

科学者の論文発表や引用における慣行に起因するその他の問題は、技術的な解決策を講じれば修正できる。第1は応用研究、とくに、科学的な分野に密接な関連のない製造工学などの分野に対する SCI の偏りである。これは部分的には SCI の取り扱い範囲が不完全だった結果である。つまり灰色文献(gray literature)と特許を除外しているのである。しかし、問題はさらに深い。装置の開発や技術の改善を直接の目標にする研究を評価、または理解するために出版物を利用することはまったく不可能だからである。いくつかの例を除いて論文間の計量文献学的分析は、国際的科学雑誌の論文が重要とされる分野において知識の進歩にいかに関与しているかを評価するのに最もよく利用できるものである。

計量文献学は出版物の種類や異分野間の引用慣行によってさらに複雑になる。図 3-1(85 頁参照)に示した種々の下位分野ごとの1論文当たり平均引用数に見られるように、引用率は分野間で大きく異なる。一般に、ウイルス学などの生物医学分野は引用数が高いが、他方で宇宙工学などの工学技術分野や確率・統計などの数学分野は平均引用数を下回る傾向がある。研究を有意義なものにするには、3つの戦略の中から1つを選ばなければならない。第1に、調査研究では同一の下位分野内の類似グループを比較することができる。言い換えれば、似た者同士を比較しなければならない。この他に、特定集団が発表するジャーナルの威信を反映するようにウェイト付けすることもできる。この「影響力(influence weight)」については後述する。他の代替策として標準

化がある。この場合には、被引用数はこれらのジャーナルにおける記事の平均被引用数と比較した相対値となる。

科学者が自分自身を引用する率が第2の区分における最後の問題である。自己引用がいろいろな形で現われて、データから得られる結論に影響を及ぼすことがないように個別事例を点検しなければならない。

計量文献学に生ずる問題の第3に関しては、科学者の論文発表の仕方を理解する必要がある。計量文献学の批判者は「過剰引用」と「過小引用」を挙げて、「不法な」引用行為についてさまざまな憶測をしている。論文がその本来の科学的価値から考えられる「引用の予想」を上回る場合、それらの引用は「過剰」と言ってもよいだろう。論文は誤りの指摘や反論のために他論文から引用されることもあるので、その場合の被引用数は、研究の質や科学への貢献以外のものの尺度になっていることもある。同様に方法や手法についての論文が科学的功績と考えられて過剰引用されるという問題もある。Lowryによれば、最も引用頻度が高くなる傾向のある論文は方法論の論文である。最後に、ある著者が本質的には同一の着想を含んでいるいくつかの短い、または反復的な論文を公表する場合、別の著者がそれら全部を引用することもあり、こうなると知的影響力のインフレーション(誇張)が発生することになる。

これとは反対の効果、すなわち、論文が「本来引用されるべき」ほどには引用されていない場合、それはにまたいくつかの理由がある。そのうち最も知られたタイプは普通「見落とし」や「無視」と呼ばれており、最初の例として Mendel の論文が挙げられる。これは出版後長い間無視されている場合である。また第2の例として Einstein の  $E = mc^2$  が挙げられる。これは広く知られているために当初論文がもはや明示的には引用されないという例である。無視による不公正の例としては数学に関する調査がある。主要な貢献をした古い論文の現在の引用頻度が、貢献度が小さい同時期の論文より低いこともあった。

論文末尾の参考文献リストに知的影響を及ぼしたもののすべてを記録しない別の理由には次のようなものがある。(1) 基本的仮定や背景知識が参考文献に引用されない、(2) ごく最近の影響が参考文献に引用されない、(3) レビュー論文の引用が少なく、着想の出所となっている個々の論文が学派によって排除されている、および(4) 非公式な影響が引用されない、などである。この他に無知、記憶違い、認識の欠如または不注意などから生じる脱落もある。

最後の批判は、被引用論文の科学的内容に関係ない引用があるという批判である。見栄えの良さは、引用を増やす。科学界の「大御所」は、古典を引用して能力を示すのと同様に、とりあえず引用される。自己利益のための引用慣行には他にもたくさん理由があり、参考文献として引用されることは必ずしも「合理的な」過程ではないこと、そしてこのため被引用数は科学としての質や重要性の指標として信頼できないものになっている。

個別の引用慣行において変則的な事象が起こる原因がさまざまに挙げられるのを見ると、計量文献学からどうしたらパターンや傾向が生じるかを考えさせられる。しかし、これらは実際に起こっており、経験ある解釈者にはしばしばわかってしまう、という事実は残る。先に報告があったように、計量文献学的指標と科学としての優秀さを示す伝統的な指標との間の関係は十分研究されてきた。ピア評価 (evaluation)、評定 (rating) または査定 (esteem)、および部門毎の順位付け (ranking) などの尺度には相関関係があることがわかっている。しかし、なぜそうなのか、それ以上に計量文献学的分析が科学者の見解と矛盾する場合どうしたらよいか、という次の質問が生じる。考慮すべき点はたくさんある。

第1に引用および出版物の調査研究の大半は、科学的成果について絶対的尺度ではなく、相対的尺度を採用している。例えば、特定分野の研究所すべてに共通な科学者行動の特徴はこれらの研究所間の比較に必ずしも影響しないだろう。無論、(構成要素への)分解の程度が高ければ高いほど、内部要因が加わる可能性はそれだけ大きくなる。第2に、引用と出版物は科学的メトリックの直接的な指標として解釈することはできない。計量文献学の指標が測定する科学的著作の特徴を記述する場合、普通混同される言葉が4つある。すなわち、質、重要性、インパクト、および引用(または出版)率である。「質」は研究がいかに良好に行われたかを示す。それは判断の問題であり、したがって異なった時期の異なった人による評価は異なるだろう。「重要性」は専門分野への論文の潜在的影響度である。科学的コミュニケーションが不完全なために、実際のインパクトはこの潜在的影響度と異なり、「影響」と呼ばれる。これらの言葉の各々は従来よりもっと不確実な、外部要因を含む。例えば、当該論文がいかに上手く書かれているか、その著者の名声はどうか、著者の習慣として、他者の論文をどのくらい読み、それらを参考文献として掲載するか否か、および当該分野の規模と成長のパターン(ダイナミクス)などである。その他の社会要因はインパクトと引用率または出版率との間に現われるが、測定できるのは引用率または出版率だけである。引用率または出版率は科学的インパクトの部分的尺度に過ぎない(この点を完全に扱っている論文: Martin & Irvine, 1983 を参照のこと)。

この方式を計量文献学に適用すれば、上述の問題の多くは解決される。引用と出版物は科学的コミュニケーションにおける最近の活動の指標であり、したがって、それらの指標がしばしば科学者自身による各自の分野の形態と分布に対する見解と一致することは驚くには値しない。しかしながら、これらの部分的指標が質と重要性の証拠になるのは、そのことが立証された場合にかぎる。このためそうした指標は他の評価方法を補うものでなければならず、その代替物ではない。

英国における経験はこうした論点を裏付ける傾向がある。計量文献学を用いる政府機関はそうした指標の主たる便益は、生産性の低い、または高い分野に速やかに注意を集中させ、この方法によってのみ明らかになる問題を提起できることであることを見出した。定量的データは「難しい」資金供与の決定を強行させ、これによりピアレビュー制度が甘くなりすぎないようにすることができる。これとは対照的にピアレビューが証拠(evidence)と矛盾する場合は、専門家には各自の立場の正当性を説明するように義務付けることができ、それによって意思決定過程をより透明にすることができる。

一般に、これらの研究資金供与機関は、例えば固定資金供与方式などに、成果指標を機械的に用いるべきではないことを知っている。むしろ個別の科学者や機関が各自の指標作成や解釈に関与する場合の方がもっと成功している。協議的手法は開放的な雰囲気を作りだし、排他的な「トップダウン」手法に必然的に伴う反感を和らげるだけでなく、研究者が自身による回顧反省の過程を通して恩恵を受けられる。

この議論の要約のために2つの表を示す。表 3-1 には引用分析に伴う共通の問題を、当該問題を最小にとどめたり、排除したりする、方法論的戦略とともに掲げてある。表 3-2 では計量文献学の長所を、各要素が評価においてなぜ重要なのかについての説明とともに示している。



表 3-1 計量文献学指標の問題点とそれを最小化する方法

Martin および Irvine の古典的表 (Research Policy, 12, p.76, 1983)の更新

<ul style="list-style-type: none"> <li>問題の分類 問題点とされているもの</li> </ul>	それは本当に問題なのか? どうしたら問題を最小にとどめ、または調査の適正な実施によって排除できるか?
<ul style="list-style-type: none"> <li>論文発表の慣行</li> </ul>	
4) 論文発表の「くせ」は機関、分野および国によって異なる	属性の一致するグループの比較や、機関、分野または国レベルでの標準化は可能である。論文発表が二次的事業である機関を評価するのに計量文献学だけを用いてはいけない。
5) 作為的操作	計量文献学指標はその他いくつかの評価手法より作為的操作がむずかしい。引用率も分析してみると作為的操作の方法の中には非生産的なものもある。例えば、短い論文をたくさん作成すればすばいほど、1 論文当たりの平均引用件数や被引用度の高い論文の点数がそれだけ少なくなる傾向がある。その他の想定される不正工作の中には合理的である場合もあるが、それは選択の影響、例えば、信望が低い論文や論文発表がむずかしく、影響度が低い分野が評価に反映される場合にかざられる。
6) 品質を示さない	引用にもとづいた尺度で、インパクトを示すものを使用すること。
<ul style="list-style-type: none"> <li>参照と引用行動</li> </ul>	
8) 下位分野により異なる	原データの集計を使用して照合グループ、例えば、1 専門分野内のグループを比較することができる。大半の評価には専門分野間の比較を可能にする標準化した引用尺度が必要である。
9) 論文貢献の種類(総説、方法論、実験、理論)によって異なる	照合グループの比較ができる。引用をインパクト指標とみなすこと。また必ずしもすべての総説、方法論の引用度が高いというわけではなく、すべての理論研究が引用されないというわけでもないことに留意すること。
10) 重要な引用	引用を品質または重要性よりむしろインパクト指標とみなすこと。
11) 自己引用	生産性の高い科学者だけが高率の自己引用を達成できる。(理由は自己引用をたくさん作成するには論文もたくさん作成しなければならぬから)。したがって、自己引用は評価結果を変えない傾向がある。出所が明確な場合は自己引用を除外することも可能である。
12) 被引用率は論文の寿命によって変化する	論文への引用は 3~4 年でピークに達する。発表後最初の 4 年間などで各論文への引用を集計すれば各論文が引用取得に同等な機会を与えられることになる。その影響が後日に明確になる「ダークホース的な」研究の影響は計量文献学や他のいかなる短期的評価によっても拾い出すことはできないだろう。
13) 特別な状況に関しては柔軟性に欠ける	著者や機関は特殊な状況、すなわち、同等の成果をあげている他の者と比較して各自の論文発表や引用集計をやりにくくなる状況に置かれるかもしれない。低得点はつねに何かに問題があることを示す兆候と見なすべきである。その理由を特定するためには、情報通の同僚に相談することができる。
14) 時間のずれ	評価を急がなければならない場合は、ジャーナルの影響力などの質を示す尺度を引用集計に替えて用いることができる。

<ul style="list-style-type: none"> <li>• <b>科学引用索引の技術的制約</b></li> </ul>	
6) 綴りの誤りや作業上の誤り	高度なアルゴリズムを用いて参照の対象となる論文と検索用文字列を照合する。
7) 筆頭著者のみの引用	包括的な著作目録を作成し、その中から論文への引用を探すことから始めなければならない。
8) ジャーナルのみをカバー	書籍がない場合は計量文献学の社会科学への適用は制約される可能性もある(もつとも、こうした傾向は変わりつつあるかもしれない)といと Hicks は見ている(Scientometrics, 1999)。計量文献学をソフトウェア開発に適用するためには会議発表論文の包括的な採録が必要になるろう。
9) ジャーナルのカバーが不十分	多くの分野、とくに基礎研究では問題ない。科学に密接に関連しない技術的分野、例えば、生産や製造などに関しではとくに問題あり。
10) ジャーナルのカバーが徐々に変化する	これが問題となる場合には、ジャーナルの固定セット (fixed subset) を用いることができる。



表 3-2 評価に用いる計量文献学主要の主たる長所

計量文献学の長所	
重要性	
<ul style="list-style-type: none"> <li>• 方法論上の長所</li> </ul>	
9) 定量的性質	操作しにくい。さらにチェックするためには「ソフト」ではない短所になりかねない問題に注目することになる。精度と一貫性。
10) 標準化の可能性	それらの周辺への影響は、中心からは見えなけれども、しばしば標準指標で見える場合がある。
11) 直接性	明確で、簡潔で、非常に単純な集計。論文作成上は何の前提もかくされていない。前提があるとすればそれは引用集計の意義に関するもので、評価の当事者全員に既知の事項である。
12) 論文発表と引用にもとづく	科学的著作やコミュニケーションには出版が必要なので、出版から生ずる尺度は評価のためだけに作成された情報より影響しにくい。
13) 時間の遅れ	引用は出版後 2～3 年でピークに達するので、引用情報は助成金授与時より 5～6 年だけ遅れる可能性がある。しかし技術的適用に関する遅れよりはるかに少ない。
14) 測定が優れている	制御不能になるまでエスカレートせず、たくさんのプロジェクト、機関などを検証することができる。各省庁から地方にいたるまでたくさんの集計レベルで同じ指標を使用することができる。
15) 成果に利害関係をもたない者が構成する分析であること	客観性が高いこと、党派の影響、同窓生の縁故関係が少ないこと。
16) 控えめ	評価対象に時間や資源を要求しない。したがって、評価の全部費用を比較すると計量文献学は比較的経済的である。
<ul style="list-style-type: none"> <li>• SCI (Science Citation Index) の技術的利点</li> </ul>	
5) 住所はすべて索引化されている。	共同研究が非常に拡大しているため、今日では論文をその他のデータベースに記載されている単一機関のみに割当てることは、その最初の著者に割当てるのと同様に適切なやり方ではない。
6) すべての科学分野をカバーしている	広範に比較できるデータベースで分野間の比較ができるようにするには標準化された指標を作り出せばよい。
7) 上位ジャーナルをカバー	科学論文は非常に広範でこれを包括的にカバーできるものは何もない。引用度が最も高いジャーナル 5,000 誌ないし 7,000 誌をカバーするという選択はカバーすべき主題の選択方法としては賢明である。
8) 引用を含む	引用分析を可能にする。

## 特許間の計量文献学

公共的研究の成果を評価する特許引用分析の利用上の方法論的問題は、これまでと同じ詳しさを論じる必要はないだろう。そこでは類似の問題が生じ、そして大部分は類似の回答を適用すると言えば十分だろう。とくに、表 3-1 の「論文発表の慣行」にもとづいて記載した3つの問題すべては、特許分析にも適用され、そして解決策は類似している。「論文参照や引用という行動」についての問題、すなわち、番号1、4、5および7などいくつかが適用される。解決策は7番-タイムラグを除いて同じである。というのも特許には、ジャーナルの影響力のように入手できる品質の代替的尺度がないからである。

## 特許・論文間の計量文献学

特許・論文間の計量文献学の威力は、結果(output)よりむしろ成果(outcome)を志向することから生ずるものである。研究の1つの成果はイノベーションの商業化に対する影響である。政府はこの種の影響に重点を置いているので、特許・論文間の計量文献学的分析に意味があることを実際に示すことが可能である。しかしながら、影響は時間の経過とともに非常に拡散し、見出しにくくなってしまっているので実証はむずかしい。したがってそれを明らかにするには経費がかかる。

米国の特許は論文を参照する頻度が次第に高くなっているため、研究と技術との関連の一部ははっきりとわかるようになった。特許は新発明とその後続く商業的用途の開発とのあいだの中間段階に位置付けられるので、これらの特許による科学的論文の引用から、技術革新と基礎的科学研究とのあいだの定量的、客観的関連を得ることができる。この尺度はいつかの用途に影響を及ぼしたが、その解釈は学界で議論にもなった。特許・論文間の計量文献学では、次の点を理解することが重要である。すなわち、1) 米国特許への引用の意味、2) 米国特許制度における情報開示の誘因、および 3) 特許・論文間の参照増加に関する解釈、である。

計量文献学者が分析した科学論文の参照は、米国特許のとびら頁の「その他の参照引用事項(Other references cited)」の部分に記載されている。これらの参照や先行特許の参照の掲載は、特許法により義務付けられている。特許法によれば、米国特許を取得するためには発明は次の3つの基準を満たさなければならない。すなわち、有益性、新規性、および当該技術に熟練した者

にとって自明な技術でないこと、である。特許のとびら頁の参照は新規性の要件から生じたものである。よって特許出願者、弁理士、および審査官は当該特許が改善しようとしている技術に対して、自分たちの知る限りの重要な先行技術をこの参照欄(references)で特定しなければならない。とびら頁の参照は特許審査官がこれを選択・選抜したものである。審査官は「入手できる一切の参照だけでなく、最良の参照を引用することが義務付けられている」(特許・商標局, 1995)。

関連指標を確信を持って使用するためには、そのベースになる特許・論文間参照がどのように作成されるのかを理解しなければならない。文書の新規性要件は上述の特許参照の一部を説明しているが、他の事情も参照の掲載行動を規定している。これらの要因の影響は、米国および欧州の特許を比較する場合にとくにはっきりしている。欧州の特許はそれほど科学論文を参照項目に掲載しておらず、参照の増加もそれほど起きていない(Narin and Olivastro, 1998)。特許に参照を付加するための手続きも欧米間では相当異なる。米国において、特許出願者は出願書類とともに先行技術の参照リストを提出するが、欧州では出願者にその義務はなく、先行技術の検索はすべて特許局の職員が行う。したがって、米国特許のとびら頁参照リスト作成には技術的内容を完全に熟知している出願者の貢献があるが、欧州ではとびら頁参照リストは特許制度を専門とするような特許局職員が作成する。

米国では、これら手続上の相違は強力な誘因によって拡大する。米国の開示規則は出願者に知っているかぎりの先行技術の参照を開示するように義務付けており、そうしない場合は特許局に対する欺瞞と見なされ、特許不許可の根拠になることがある。欧州に比べ米国では特許訴訟のリスクが大きくなるため、参照の掲載がますます必要になってくる。審査官は参照文献を審査し、そして当該発明が文書に報告されている技術に比べて斬新であると決定したものと見なされる。したがって特許に参照されている文献は、裁判所にとって、「当該特許は発明が斬新でないため無効である」という証拠として利用するのが難しくなっている。

米国制度では先行技術開示という強力な誘因に促されて参照を特許に添付する手続きを経るため、特許・論文間の関係が明確になる。これをさらに追求すると、審査官や出願者は不適切な資料を山ほど参照(引用)してしまうようになるのではないかと考えてしまう。しかし、こうしたことは起きないようである。強力な誘因が働いたとしても、それは関連する先行技術の場合に限られるからである。不適切な先行技術への参照を探して添付するために時間や資金を費やしても、法的保護を強めることにはならない。科学と技術のあいだの関係を特徴付けることを望む分析者は、米国

特許制度が関連する非特許先行技術を文書化する強いインセンティブを与えていることには満足している。特許における「他の参照」に科学論文の参照を見つけることができるからである。

特許参照の記載をめぐる特許局の手続きと法的要件は、特許を取得した技術に関連する科学を示すものとして指標を解釈する上での正当な根拠を提供する。さらにこうした解釈を裏付けるのは、専門家も確認しているように、科学と関連の強い技術は科学を参考文献として記載することが多いというパタン存在である。指標の解釈は参照を特許に添付する社会的過程の特定のモデルを基盤とするわけではないことに注意すべきである。例えば、参照記載行動のかなり機械的なモデルは次のようなものかもしれない。すなわち、発明家が論文を読み、発明に関する着想を得て、論文を参照する特許出願書を提出する、などというものである。このモデルがまったく妥当でないらしいことは、Meyer が行ったノルウェーとドイツにおける 10 件の特許の検証が示している (Meyer, 2000)。この他に機械的モデルが妥当でないことを示しているものとして、科学的参照記載に関する研究があげられる。こうしたやり方では、科学的参照記載の説明のために提案された類似の合理的メカニズムに対して経験的な裏付けを見出すことはできないだろう。参照記載の合理的モデルは、科学界では有望のように見えたとしても支持は受けられなかった。結局、著者が各自の論文に参照を添付することと、米国の特許申請書の場合とは対照的である。米国において、特許申請者は審査官にとびら頁に参照を記載してある旨提示するに過ぎない。

研究と技術のあいだの密接な関連が特許による論文の引用のパタンになって表れる社会的過程が、法的な制約を受けて多様かつ複雑になることは明らかである。科学集約的技術の場合、発明家は論文を書き、それによって引用対象の出版物を得るだけでなく、その特許に参照される論文を増やすような研究論文を熟知できるように見えるかもしれない。審査官は博士号の資格を持つほうがふさわしく見えるかもしれない。博士号は科学集約的特許の検証に必要な資格だからである。寄与要因をすべて引き出すには、特許構造の民族誌的研究が必要になるだろう。しかしながら、技術の科学集約度は当該技術の特許が科学論文を参照する率に反映されるとする解釈は、特定の社会的過程(合理的かどうかを問わず)の運用にもとづくものではない。

特許・論文間の引用指標の解釈は複雑である。複雑な社会的過程や特許局の規則が特許記述の根底にあるからである。特許・論文間の参照記載件数が増加している事実の解釈はさらに複

雑である。社会的過程や特許局の規則が、科学と技術のあいだの変化よりむしろ指標に作用することがあり得るからである。

CHI は、1987/1988 年と 1993/1994 年とのあいだに特許による論文の参照が 3 倍になったことを見出した(Narin et al., 1997)。Narin はこれをイノベーションの新しいパラダイムの証拠だと解釈した。このような枠組み(パラダイム)では新興技術は研究をイノベーションに直接転換することによって進歩するが、これはもっと伝統的なモデルと対照的である。伝統的モデルでは技術は主として先行技術の進歩にもとづいて構築され、そして研究への結びつきは間接的である。この顕著な結果は計量文献学者やエコノミストにこう言わしめた。「イノベーションに変化が生じないのに、特許による論文の引用が増加したのは特許取得過程が変わったためかもしれない」と。これには 3つの要因が挙げられた。開示規則の強化、特許訴訟件数の増加、およびデータベースへのアクセスが容易になったことである。

これらの仮定を検証するために CHI は参照記載件数が増加するパターンを調査した(Hicks et al., 2000)。上記の3つの要因は異なった技術にも異なったタイプの参照にも別々に影響することはないので、これらの仮定による予測では、増加率は参照のタイプ(米国特許、他国特許および科学ジャーナルの論文)や全技術のどれをとっても等しいだろう。増加率が技術のあいだで、または特許における参照と論文におけるものとのあいだで異なる場合は、上記のようなおおまかな説明は適切ではない。CHI により、科学論文を参照として掲載する件数の増加は、米国特許へのそれをはるかに上回っていることがわかった。これは開示規則の強化が科学論文の参照件数の増加のすべてを説明できるとは限らないことを示している。米国特許への参照記載件数の増加は科学論文の場合に比べて、それほど大きな技術による違いがないが、この意味は、大まかな要因ではなるほど米国特許への参照記載の増加を説明できるが、科学論文を参照として記載する件数の増加は説明できない、ということである。

論文データベースへのアクセスが容易になると、論文への参照の記載件数の増加を促すという考え方は別のレベルでも信じがたい。注目すべきは、出願者がそれぞれの分野の専門家と同様に各自の発明に関する論文を持っている場合は、データベースへのアクセスのしやすさが彼らにはそれほど重要ではないことである。米国特許審査官にとってデータベースへのアクセスのしやすさが 1980 年代から 1990 年代にかけて増加したとは考えられない。科学論文を探すダイアログのユーザーたちは 1972 年以来強力な検索が可能だった。しかしながら、ダイアログの使用法は

複雑であったし、近年ではもっと簡単な検索方法が見つかって、そうした複雑な操作を学ぶ必要は少なくなった。しかし、このことは、非常に狭い「技術」を担当し、先行技術の発見に多くの時間を費やしている米国の特許審査官に、どういう関係があるのか？こうした状況で、文献の出所へのアクセスにダイアログを利用しなければならなかった者は非常に頻繁にそうする必要があるのであろうし、またその方法の習得のために必要な時間を投資したとしても当然である。さらに、特許局は審査官に必要なだけの情報源を提供しようとしている。例えば、1995年まで外国特許が重要な分野では外国特許を入手して米国流に分類し、審査官の利用に供していた。よって、データベースへのアクセスが容易になったという事実は、科学参照記載件数の増加を説明しているわけではないようである。

結局、特許から論文への参照記載を科学と技術の関係の指標として開発するために、ほぼ20年間研究が行われてきた(Narin, 1976; Carpenter & Narin, 1978 and 1983; Carpenter et al., 1980; Carpenter, 1983; ABRC, 1986; Collins & Wyatt, 1988; Narin & Olivastro, 1992 and 1998; Narin et al., 1997, 1998)。明らかにこれらの指標は多くの手段のうちの1つでしかない。とは言え、それらは特別な長所をもった貴重なツールであり、これにより技術、国家および時空間にわたって、科学・技術関係の定量的比較ができるのである。

## 方法の要素

### 論文と特許の集合の定義

計量文献学による評価の第1段階は、検証すべき論文または特許の集合を定義することである。これには根本的に異なる2つの方法があり、どちらを選ぶかは当の研究の性格による。第1の方法は著者、機関またはその他の分析項目について書誌を編纂することである。第2の方法はテーマ別の書誌を編纂することである。

著者、機関または国の書誌は、データベースがいかに整っているか、およびソースが入手できるかどうかによって、多かれ少かれ難しい問題がある。著者目録の場合は誤表記と思われる論文を探さなければならず、また同じ名前の他人の著になる論文を排除しなければならない。これは時間のかかる仕事だが、必要な第1歩である。CV または他の書誌は出発点としては申し分ないようだが、この2つを同時に扱うのは非常にむずかしい(Dietz et al., 2000)。また、それらに

は誤りや欠落情報が含まれ、一貫していない。言い換えれば、人または機関の方がまだ完全に近いということである。このため補完的検索が必要であり、書誌からは、このような論文検索は不可能である。

テーマ別の書誌につきまとう問題はさまざまである。あるテーマの特許または論文における書誌作成の手段は、フィルターまたはプロトコル(手順)の開発である。このプロトコルは、関心テーマの一組の論文または特許をもたらす反復実験により見出だされるような、文献や分類や引用関係において、住所の組み合わせ、単語および単語の組合せを規定する。各セットは関係のないものを取り除く必要がある。人工的に作り出した変形はフィルター要素のいくつかを使って文献を集めるアルゴリズム形式での作業に適用されるが、実際の情報を使ってこうした過程を通して作業することも確実に可能である。無関係な資料を取り除いて完全な文献集合をそろえることが評価には重要なので、アルゴリズムの結果は常に点検して完全性や精度を確保すべきである。

### 全体または案分の集計

著者、機関または資金供与源はそれぞれ、論文の元になっている研究にどれくらい貢献してきたか？ 理想は、当事者全員がそれぞれの貢献率について合意に達することであるが、現実には、関係者のあいだに合意はない。各自が自己の貢献度を自分たちの協力者の貢献よりも少し重要なものとして、格付けする可能性が高いからである。機関はインフラストラクチャや設備の提供を考慮に入れるので、おそらく著者には同意しないだろう。そして資金供与機関も同様に妥協の余地のない異論を唱え、協力者を研究室に呼ぶための助成金の方が博士号の学生が受けた支援や器具を給付した助成金より重要であると主張するだろう。

計量文献学の世界では、論文にクレジット(貢献の認定)を付与するのに2つの適切な選択がある。案分集計と単純集計である。(この他、筆頭著者、機関または資金供与源などにクレジットを付与することもあるが、これはおかしいやり方である)。案分集計では、著者・機関・資金供与源、それぞれは他にどれだけ掲載しているか、その比率によってクレジットを付与される。そこで、著者が2人いる場合は、各自が付与されるクレジットは論文の半分についてである。単純集計では、各自が関与する各論文に対して丸ごと1単位のクレジットが付与される。

案分集計には計算上の長所があり、加算合計が 100%になって、二重集計はない。単純集計される場合は、論文の点数は集計完了後に著者・機関・資金供与源ごとの加算をすることができない。言い換えれば、いかなる集計も共同研究の二重集計を避けるためももとの計算にあらかじめ織り込まなければならない。

案分集計には共同研究を不利に扱うという短所がある。最近では共同研究は一般に善と考えられ、そして各政府がこれを称揚しているので、どんな評価方式でも共同研究を理由にペナルティが設けられている場合は、その方式は疑わなければならない。共同研究をすればそれだけ皆の仕事量は減るべきだし、共同研究によって少なくとも2倍の引用を得ることができなければ、協力すべきではない。したがって案分はペナルティの対象ではないという経済学者の見解を取ることでもできよう。あるいは、協力するのは手抜きをするためではなく、協力しなければ不可能な研究を実現するためであるとも考えられよう。案分集計によって課せられるペナルティを克服するほど十分ではないが、協力による論文は著者が単独の場合より引用度が高いという事実もある(Katz, 1997)。2人の著者が別個に仕事をし、そして例えば、5回引用される論文を2点作成するのと、協力して 10 回引用される論文を1点作成するのと、どちらがよいかを冷静に決定しなければならない。著者たちは分担によっていずれの方法でも同じクレジットを得る。知識への貢献度は共同研究の方が甚大だとしても。

## 機関の特定

一般に、論文は評価に使えるようにするためにはその機関を特定しなければならない。著者の住所はすべて *Science Citation Index* (SCI) で特定されているが、これは評価の観点からすれば、SCI 独自の長所である。残念ながら住所は機関の関連事項ではなく、各論文をその所属機関で特定するためにはさまざまな表記の住所を統一的に扱う辞典(シソーラス)が必要である。こうしたシソーラスは評価インフラストラクチャの貴重な一部になる。シソーラスがなければデータは各調査研究ごとに処理しなければならない。

## 資金供与源

CHI 社の評価研究調査の多くは、政府機関が資金供与した論文の分析を含む。この情報はデータベースには含まれないが、調査のつど論文レベルで CHI 社が編纂する。この手続きに便法



はない。スタッフは図書館で論文を探し、著者が謝辞に記述した資金供給源を記録する。論文の何%かは見つけられない場合がある。古い論文は新しい論文より情報が欠落していることが多いようだ(したがって、特許に引用されている論文について資金供与源を確認するという CHI 社のほとんど日課になっている作業では、11 年の分析期間が使われている。これは当該特許発行の 11 年前まで遡って発行された論文を点検しているということである)。また、必ずしもすべての著者が資金供与について記述しているわけではない。部内に資金供与源をもつ会社や研究所に所属している著者は、資金供与に対する謝辞を述べないと考えられるが、彼らの資金源は彼らの所属から推定できる。さらに、謝辞の記述はジャーナルや分野によって異なる。レビュー論文では資金源を記述する 경우가少なく、これらの論文の引用頻度が非常に高い場合は、資金源確認を検証する研究を注意して行わなければならない。

## 被引用回数の集計

直接引用手法の場合は、一組の研究論文の影響度や質は論文刊行後最初の3年から5年、またはそれ以上のあいだに論文に対して行われた引用回数を集計することによって行われる。このためには引用情報を含む唯一の大型データベースである *Science Citation Index* (SCI) を使用しなければならない。また各引用はこれを参照した研究論文と照合しなければならない。直接引用集計手法を用いると、論文が引用されるまでに多年の経過を要し、また、標準化すると図 3-1 に示すような下位分野ごとに引用率が違うことを説明できる。直接引用集計手法を使えば、影響力の大きい個人・トピックス・研究グループなどを特定することができる。下位分野と年別の引用分布を知っていれば、一群の論文への引用を下位分野や引用年と被引用年の異なった組合せに対する平均値と比較して高、低、または平均として特徴付けることができる。

計量文献学の長所と短所に関する方法論的議論の結果、厳密に釣り合いの取れているグループを比較しないかぎりには、標準化した引用尺度を使用しなければならない。つまり、論文の出版年、ジャーナル、分野、国などを所与とすれば、引用集計の高、中、低を評価する作業には基準点(ベンチマーク)を設定しなければならない。こうした基準点の設定方法はいくつかある。

一般的に利用されている1つの方法は、ジャーナル単位で標準化することである。つまり、同じ年に同じジャーナルで刊行された論文が受けた引用の平均回数にもとづいて各論文の見込み被引用率を設定することである。この基準を使って、ある日まで論文ごとに累計した被引用回数は、

当該ジャーナルの当該刊行年における論文に対する見込み被引用回数と比較することができる。この方法の問題点は、被引用回数が累計で5件だった弱小ジャーナルの論文が、被引用回数が累計で10件にのぼった有力ジャーナルの論文より、重要と見なされることがあるという点である。研究者には *Nature* や *Science* における刊行を避ける誘因となる。そして、ある機関が発行するジャーナルの質を改善する場合、その機関のパフォーマンス指標は下落するだろう。この種のナンセンスを避けるために、標準化には基盤を広くとることが望ましい。

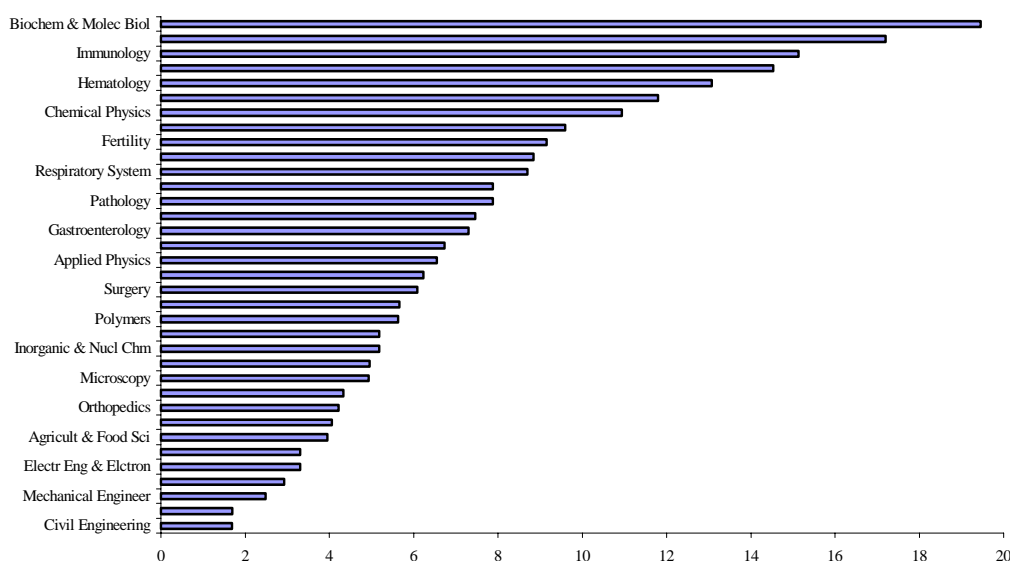


図 3-1 Citations per paper vary by subfield

Average number of citations earned by 1989 papers in the 5 years following publication using a 1997 fixed journal set

CHI 社は下位分野にもとづく標準化を利用しており、そしてその実施にはいくつかの方法が考えられる。最も高度な、最も初期の方法では、統計理論を使って標準化した引用得点を導いた。標準化引用得点の論拠は、受けた引用の回数を使って論文を比較するよりむしろ、同じ分野の全論文の引用分布におけるそれらの相対的地位を使って論文を比較することであった (McAllister et al., 1983)<sup>41</sup>。この方法は評価に利用されたが、その後はやや利用されなくなってきた。というのも、引用回数の集計と得点のあいだには統計的計算が媒介するため、得られた結果の直接性やわかりやすさが減じられるからである。

<sup>41</sup>厳密に言えば、同一分野における米国論文全体の分布が利用された。米国の論文が評価されていたからであった。無論、その他の全米規模の分布や特定の国から選んだ文献の分布も比較基準として選択することができるだろう。

もっと頻繁に利用されるのは、評価対象論文1点当たりの平均被引用回数と下位分野全体に対する論文 1 点あたりの平均被引用回数とのあいだの直接比較であった。ここでの引用回数は刊行後一定期間にわたって集計された。この一定期間を「ウィンドウ」と呼ぶ。通常4年から5年のウィンドウが使用された。たびたび使用される変形方式では、その下位分野の上位 10%、すなわち、上位十分位数の中に見出される評価対象論文のシェアが報告された。上位十分位数の中に含まれる論文はとて多く引用されており、評価対象論文のどのような組合せにおいても、それらの 10%は最も引用度の高い下位分野の 10%であると見込めるだろう。したがって、もし論文の 20%または 30%が下位分野の上位 10%に含まれていたら、これは非常に大きな成果といえる。次章においてこれらの標準化手法の例を挙げる。

ジャーナルにおける下位分野の分類は、標準化に関しては根拠が余りに不完全であり、したがってより細かい区分け、例えば、「環境科学」の代わりに「水中汚染輸送の研究」というように区分けを行うべきであると論ずることもできる。ある意味ではジャーナルは標準化のためのより細かい区分になることなど、我々は本章ですでに問題の一部を見てきたし、そうした方法の危険にも触れてきた。もっと細かい区分けも「金魚鉢の問題」とでも呼ぶべき問題に左右される。政府機関が次の質問をしても問題ないだろう。「我々が研究者に達成するように要求している基準とは何か?」。基準には国際、国内または地方の3つのレベルがあるかも知れない。評価が取り組むべき問題は「取った魚はどれくらいの大きさか?」一標準を設定すると、これは次の意味になる。すなわち、「海洋、湖沼、水族館のうちどこにいる最大の魚を欲しいのか?」。水溜まりは魚が大きく見えるように小さくすることもできるが、獲物は金魚鉢の中で最も大きい魚であるとわかれば情報としては十分ではないか? 下位分野レベルでの標準化は合理的な選択である。数学者が不可能を達成したり、生物医学研究に匹敵する被引用率に到達したりすることを要求していない。他の4点の最も関係深い論文と比較した上で、または一般に低品質のジャーナルに掲載されたからという理由で、どんな隠れた数学論文でも立派であるとは保証していない。

## ジャーナルの影響度

引用集計は各論文についてまとめられるが、ジャーナル影響度はジャーナルをもとにした尺度である。ジャーナル影響度を用いる目的は、引用集計のように妥当な引用歴を蓄積するのに4年も5年も要しない、すぐに得られる代役となることにある。ジャーナル影響度というアプローチは、1

組の論文の影響を測定するために正確な引用集計をせずに、各ジャーナルの得点を利用する。ジャーナル影響度による得点は本質的には各ジャーナルの論文が引用される回数の加重平均であり、したがって、著名で引用度の高いジャーナルにおける発表は著名度が低いジャーナルより影響度が大きいと見なされる (Pinski & Narin, 1976)。影響度計算においては各ジャーナルの当初影響度を1と仮定する。次いでジャーナルが他のジャーナルで引用される回数をジャーナルが行う参照記載回数で除される。これはジャーナル影響度への一次近似値、所与の参照に対する被引用の比率でもある。次いで最初の反復の終わりにこの影響度は当該ジャーナルからの参照件数の加重計算の一部として利用される。すなわち、引用度の高いジャーナルからの引用は普通のジャーナルからの引用より重要であると見なされる。反復は各ジャーナルに対する安定した影響度が得られるまで続けられる。その影響度は、1論文当たりの影響度を割り出すために、ジャーナルに掲載された論文の数で除す。最終結果は、1論文当たりの影響度は、本質的にジャーナルの平均的論文が受ける1論文当たりの引用の加重回数になる。20点から50点あるいはそれ以上の論文の大型組合せにおいては、1論文当たりの影響力は論文の長期引用実績の適切な近似値になる。もっと一般的に使用されるジャーナルの尺度はISIが計算したジャーナルのインパクトファクターである。これは各ジャーナルについて、影響尺度の反復計算を加えずに1論文当たりの平均引用回数を(少し緻密に)測定する(影響要因を評価に使用する点についての警告に関する報告、Seglen, 1997を参照のこと)。

ジャーナル影響度を測る大きな長所は、それが簡単に使用できるのとすぐに利用できる点にある。個々の論文が何回引用されるかを知るのに3年から5年も待つには及ばない。こうしたタイミングの良さはあるが、その代わり、特別に引用度の高い科学者や科学者グループ、引用度の高い論文の特定はできず、かつ論文数が多くなると統計的検定が必要になる。

最近、影響力方法論は、コンピュータ科学者によってウェブページの格付けをするように改造されている (Kleinberg, 1999)。彼らの影響度を組み込んだウェブページの格付け方法は、グーグル検索エンジン([www.google.com](http://www.google.com))に組み込まれた。

## 分野と下位分野

本章で報告したCHI研究で使用されている第2のジャーナルレベルの要素は、研究を刊行したジャーナルにもとづいて科学の下位分野と分野に論文を分類することである。CHIのジャーナル

分類システムは、SCI の各ジャーナルを約 150 の下位分野の1つに当てはめる。これらの下位分野は9つの分野に統合される。大部分のジャーナルは 1 つの下位分野にうまく適合するが、*Science* や *Nature* をはじめとする約 100 点のジャーナルは、範囲が非常に広範にわたっているため一般分野に含められる。これらのジャーナルに載った論文を、ジャーナルが引用する分野にもとづいて個別に分類するという方法は、最も進歩した扱い方である(Glanzel, 1999)。

## 研究レベル

CHI 研究において使用しているジャーナルレベルの第3の分類パラメータは、研究レベルである。研究レベルの背景にある着想は、応用研究範囲にとって基礎となるある機関の出版を特徴づける単一の数値の開発であった。応用(研究)ジャーナルがより基本的なジャーナルを引用する傾向があるという観察からはインスピレーションが沸いたが、その逆は真ではなかった。つまり基本的ジャーナルは応用ジャーナルを引用しなかった。この方式を実行するため SCI が取り扱うジャーナルはそれぞれ、CHI 職員により4つの研究レベル、つまり「レベル1」の応用度が高い技術から「レベル4」の最も基礎的な研究にいたるレベルに分類された。表 3-3 は4つの研究レベルのそれぞれに対する代表的ジャーナルを示して当該方式を説明している。注目すべきは、*Physical Review*などの非常に基礎的な物理学ジャーナルはレベル4であるが、*Journal of Applied Physics* はレベル3であり、典型的なエンジニアリングジャーナルはレベル2、またレベル1は研究ジャーナルよりむしろ非常に応用的な専門誌だということである。

分野の相違は各レベル内の論文の比率から生ずる。物理学はほとんどレベル1でも2でもないが、エンジニアリングの下位分野はほとんどがレベル1およびレベル2である。いくつかの下位分野、とくにがん研究などの生物医学の下位分野の一部では、非常に臨床医学的なものから、非常に基礎的なものまで多様なジャーナルがある。

表 3-3 研究レベルとそれらのプロトタイプ・ジャーナル

レベル	物理学および生命科学における名称	プロトタイプ・ジャーナル
1	応用技術	日本鉄鋼協会『鉄と鋼』
	臨床観察	Gastrointestinal Endoscopy
2	エンジニアリング・技術科学	J Nuclear Science & Technology
	臨床ミックス	New England Journal of Medicine
3	応用研究 Applied Research	J Applied Physics D –Applied Physics
	臨床研究	Cancer Research
		J Clinical Investigation
4	基礎科学研究	Physical Review B –Condensed Matter
		J American Chemical Society
		J Biological Chemistry

## 結論

この数十年間 Francis Narin と CHI リサーチは計量文献学指標の可能性を完全に生かすための開発を行ってきた。開発された関連指標は、論文間、特許間、および特許・論文間の3種である。これらそれぞれに長所と短所があり、解釈には注意が必要である。指標の解釈にも注意が必要である。指標のもとになっているデータベースは、そのままの形式では分析作業に適さない。分析的研究を行うには、シソーラス、分類および基準が必要となる。次章では研究評価に関する質問に回答するため、これらの分析的要素を利用する計量文献学が実際にどのように展開しているかを説明する。



## 第4章 計量文献学を利用した評価研究の例

報告者: Diana Hicks, Peter Kroll & Patrick Thomas

本章は計量文献学を利用し、かつ政府機関がそのプログラムの検証のために委託した研究の、短い要約で構成されている<sup>42</sup>。それらは4部からなっている。第1部では CHI が行ったものではないが、計量文献学を含む最近の評価研究を説明している。さらに、第1部には Bozeman et al.による最近の定量的事例研究(研究価値マッピング・プロジェクト)の説明も含まれている。この章のその他の部分では米国連邦政府のために 20 年間にわたって CHI が行った作業を説明しており、論文間、特許間、特許・論文間の計量文献学という、3部に分けられる。要約は各節の中で年代順に配置してある。これらの事例研究は大半が CHI リサーチ社の資料であり、そしてある場合には作業を委託した政府機関の身元は削除してある。本章の目的は、実施された計量文献学研究の領域や、計量文献学の種々の利用法、この手法によって提起される疑問の数々、を説明することである。

第3章に説明した計量文献学方法論は、第3章で説明された「ジャーナル影響度」または「標準化」などの事柄に触れるこれらの要約の、参考になる。さらに、本章で説明する研究は第1章の裏付け資料になる。第1章では当時非常に広範に知られた NIH のための初期の研究を説明し、成果(outcome)に関連する研究の最近の発展もまた特許間や特許・論文間の研究に見ることができ

---

<sup>42</sup> 理想的には各要約は、研究の受入れや利用に関する注解を含むべきであったと思われるが、ありがちなことに、そうした情報は大半の事例では知らされていない。



## 事例の一覧

A. 米国連邦政府のために行った定量的データを含む最近の研究評価の例	p. 95
1. イノベーションに対する米国ファンディング機関の影響の評価	96
2. 資金を供与された基礎研究の「成果」の評価	100
3. 技術分野の重要インフラストラクチャ特定のためのテキスト・マイニング	102
4. 米国研究分野のベンチマーキング	105
B. 論文間の分析: 研究結果(output)の評価	109
1. 研究目標の促進を求める法律の効果の評価	110
2. 研究所の出版物の包括的なベンチマーキング	112
3. 資金供与メカニズムの比較: センター・プログラム方式とピアレビュー方式による資金供与	115
4. 研究機関における日常業務としての計量文献学的プロファイリングの例	118
5. 重要ながん研究上の発見に影響する要因の評価	124
6. 内部的または外部的な研究支援メカニズムの比較	127
7. 助成金受領者の成果と助成金申請が不採択だった者の成果の比較	129
8. 技術移転プログラムが科学者の論文発表に及ぼす影響の評価	131
9. 世界的な競争者に対する研究所の発表論文のベンチマーキング	133
10. 伝統的にマイノリティの多い研究機関の研究能力向上のためのプログラム、RCMIの影響の評価	136
11. 科学および社会科学研究を結びつける学際研究機関のアウトプットの評価	138
12. 計量文献学をウェブに拡張して成果(outcome)を評価することができるか?	140
C. 特許間の分析: 研究成果(outcome)の評価	143
1. 技術移転プログラムの科学者の特許取得に対する影響の評価	144
2. 共同発明者クラスターの知的財産権管理への応用	146
3. 研究所の特許ポートフォリオの評価	148
4. 特許ポートフォリオの引用分析による技術移転の評価と増加	150
5. 連邦機関が維持管理しているデータベースの研究および技術に対する貢献度の評価	152
6. 政府特許ポートフォリオのマネジメントのための特許引用マッピング	154

D. 特許・論文間の分析:研究と技術の連関(linkage)による成果(outcome)の評価	157
1. 科学ベースの技術において、一国の役割は応用が進むにつれて消滅する。	158
2. ある研究所の類似研究所に対する特許権取得のベンチマーク	161
3. 特許に引用された研究所論文の検証 – 3種の報告書から	163
4. 政府機関の科学研究を参照する特許技術の評価	166
5. 眼科技術に対する NEI 資金供与研究の重要性の検証	169
6. 特許に引用された政府機関資金による論文の特徴の検証	172
7. 最近承認された医薬品における NIH が資金供与した研究の重要性の検証	174
8. 中核分野外における政府機関支援科学研究と特許技術とのあいだの関連性	177
9. 科学的基盤を NEI 資金供与研究にたどれる技術の経済的影響の検証	179
10. 科学的基盤を NEI 資金供与研究にたどれる派生的技術の経済的影響の検証	183
11. STTR 技術に対する公的資金供与研究の影響の検証	186
12. 特許に引用された論文に対するオーストラリア政府の資金供与の特徴の検証	189



A.

**米国連邦政府のために行った  
定量的データを含む最近の研究評価の例**

## 1. イノベーションに対する米国ファンディング機関の影響の評価

Roessner, D., B. Bozeman, I. Feller, C. Hill, N. Newman,

*The Role of NSF's Support of Engineering in Enabling Technological Innovation*, first year final report for National Science Foundation (SRI International: Arlington VA, January 1997).

Available at: <http://www.sri.com/policy/stp/techin/>

Roessner, D., R. Carr, I. Feller, M. McGeary, N. Newman,

*The Role of NSF's Support of Engineering in Enabling Technological Innovation: Phase II*, final report to National Science Foundation (SRI International: Arlington VA, May 1998).

Available at: <http://www.sri.com/policy/stp/techin2/>

### 目的

工学、とくに研究およびその関連活動に対する NSF の支援が最近の重要なイノベーションの開発と商業化にいかに関与したかを調査すること。

NSF が挙げている目標は以下の通り:

次の目的のためにこれまでの 10 年間に出現した工学上の最も重要な 12 件のイノベーションを促した過去の発見、出来事、人々、相互作用および条件などの体系的な検証を行うこと。(1) イノベーションを生むに際しての NSF の関与を記録すること、(2) イノベーションの開発のより広範な背景における NSF の役割の大きさを評価すること。

### 方法論

当該研究に当たっては、科学研究の技術に対する影響の追跡を回顧的事例研究で行う方法を用いた。この方法の先鞭をつけたのは 1960 年代後半から 1970 年代全般にかけての、かの有名な研究、ヒンドサイト・プロジェクトおよび TRACE プロジェクトであった。この研究は科学における重要なイノベーションの起源の特定を目指した研究であった。研究の対象になったイノベーション

とは次のものであった。インターネット、磁気共鳴画像 (MRI)、反動射出成形 (RIM)、コンピュータ援用設計の電子回路への応用 (CAD/EC)、電気通信用光ファイバーおよびアナログ携帯電話。

### 関連方法

1. 各イノベーションを支える技術の特定、およびイノベーションの本質的な部分と既存の技術を支援している部分は何かを決定すること。
2. 技術開発を記述する研究のすべてを探索するためのオンライン・データベースのライブラリー検索。
3. 計量文献学—計量文献学における NMR (核磁気共鳴) の非常に小規模な、試験的実験が ISI の研究活動分野データベースを使って行われた。これらのデータベースは出版物のグループ分けにクラスタリング手法を用いている。データ検索の目的は、NMR の分光研究に対する NMR 画像との密接な関係を示し、かつ NSF が資金を供与している研究者を複合材料分野における主要貢献者の中から探し出すことであった。特許データベース検索の目的は発明家、機関、共同発明者および主要研究論文の引用などを探ることだった。2回にわたる追加実験では論文と特許における共通引用分析方法を用いた。
4. 機関分析—ライブラリー検索、NSF スタッフとの議論、インタビューおよび NSF 授賞者のデータベースの検索などから技術を開発した主要組織を特定した。
5. 個人および電話インタビュー—上記手法を用いて特定された貢献者に当該技術と NSF の役割についてインタビューを行った。

### 結果

6件の事例のほぼすべてにおいて、政府とくに国防省部門による研究技術開発への支援は大きな役割を果たした。それらの事例から例外なく明らかになったのは、教育訓練、とくに大学院教育などが工学イノベーションに本質的な役割を果たしたことだった。実際、上記6事例すべてにわたって目立った単一の、一貫したパターンがあるとするれば、それは個人発明家、技術企業家および継続的に技術進歩を促す可能性がある最新技術で教育訓練を受けた学生などの人的資本が演じる重要な役割だった。規制政策が携帯電話や反動射出成形事例におけるイノベーションの過程を形成した。優れたイノベーションが一般的にただの概念から市場における成功に発展

するのに数十年かかるとすれば、基礎研究が上記の工学イノベーション6事例において中心的な役割よりむしろ支援的役割を果たしたことがわかったのは驚くことではない。

NSF は各イノベーションにつねに大きな貢献をしている博士号クラスの科学者やエンジニアの教育訓練の大きな、時には主たる支援元の役割を一貫して果たした。NSF の大学研究インフラストラクチャの支援非常に有力な活動の第 2 である。NSF はインターネット、CAD/EC および MRI など、半分の事例においてインフラストラクチャに大きな支援の手を差し伸べた。その直接の研究支援は、1 つの事例、CAD/EC では成功したイノベーションのかぎになった。その研究支援はあらゆる事例の推移に不可欠な知識を育んだ。その組織的リーダーシップはインターネット事例では卓抜しており、顕著でユニークだった。約半分の事例では、NSF は産学の研究者を有望分野の議論に結集させるワークショップやシンポジウムを通じて、大学研究者に産業界の問題に取り組むよう促し、研究の新しい流れを作った。

#### 計量文献学の結果

RIM-294 特許は「反動射出成形」なるキーワードを使って発見された。64 件の特許は5回以上引用された。7回以上引用された特許は実質上すべて RIM に関連したもので、民間会社に発行された。科学論文を参照した特許の比率から、当該分野は Narin の論文データと比べた場合、科学に関連付けられていることを示している。

#### NMR

ISI の *Research Front Database* が用いられたのは、1990 年版 *Research Front Database* において専門分野クラスターのマップを作成するためだった。マップは画像と分光学それぞれに関連した2組の相関する専門分野を明らかにした。2分野間の強い関係がインタビュー結果から確認された。この実験から得られた1つの仮の結論は、計量文献学によるマップ制作手法はイノベーションに関する研究テーマの境界を作るための役に立ちそうなツールを示している。

#### 携帯電話通信技術－論文

1988 年、1989 年および 1990 年の *Research Front Database* には著者が引用するか、引用されるケースとして5つの名前が見つかった。*Research Front Database* における論文数が集計された。論文関連の 46 のグループの内の1つは当該分野の先駆者別の論文を含んでいたが、それ

らはこの種の検索が事例研究の開始に際して一連の知識の有効性を証明し、提供する貴重な方法であることがわかるだろう。

#### 携帯電話通信技術－特許

オンライン検索では引用度の高い 10 件の特許が特定された。これら 10 件やその他などで引用された特許が検索された。これらの特許の共通引用分析が行われ、マップが作成された。そのマップにはすべての多次元尺度構成によるマップによくあるように歪みがあった(2 つの最も密接に関連している特許はマップでは関連が薄かった)。そのマップによって AT&T とモトローラが当該技術開発のプレーヤーであるという結果が確認された。マップは大まかで歪曲があるが、デジタル技術の到来を暗示しているだけでなく、携帯電話通信技術の基礎になっている特許技術を概観するのに便利である。当該手法は開発を進める価値があり、また事例研究に関する技術の理解に貢献することができる。

#### コメント

事例によって豊富な詳細や意思決定やさらなる研究を深めるのに役に立つ技術イノベーションを可能にする上での NSF の役割についてニュアンスを含んだ理解が得られる。

計量文献学はそれほど成功したわけではなく、フィルターが得た成果だけにとどまったようだ(方法論の章を参照のこと)。フィルターまたはマップによって研究分野を区分けする段階は理論上は最終段階ではなく最初の段階と言えるだろう。標準化が行われていないため、作成データの解釈はできなかった。



## 2. 資金を供与された基礎研究の「成果」の評価

Barry Bozeman et al. - aka: The Research Value Mapping Project, "Qualitative-Quantitative Case Studies of Research Projects Funded by the Office of Basic Energy Sciences," Final report submitted to the Office of Basic Energy Science, Department of Energy, February 1999.

Available at: <http://rvm.pp.gatech.edu/papers/index.html>.

### 目的

エネルギー省、基礎エネルギー科学局 (BES) のプロジェクトの評価

### 方法論

4年間にわたって BES が資金供与したプロジェクトの 28 の事例研究と1件の調査が行われた。研究から成果にいたる知識の流れを追跡する分析的モデルが開発された。成果は分岐構造における一連の事象としてモデル化された。偶然の要因を成果に結びつける仮説が開発された。データはインタビューと文書から収集された。次の事項のための指標が開発された。資金供与額などの費用便益変数、人数、教育訓練を受けた人の数、収入額、技術移転事務所の存在、ユーザーが新製造方法などの開発を必要としたかどうかなど。定量的分析はどの変数が成功と関連したかの決定に利用された。

### 結果

BES 資金供与が安定していることは大事なことである。BES 資金供与はしばしばその支援を受けている研究者によって中核資金と見られ、またこの資金によって研究者が重要な業績を上げる能力を構築し、知識に貢献し、学生の教育訓練ができるようになった。

新たな組織設計はプロジェクトの効率性を向上させる。BES プロジェクトは多様な革新的な機関・組織設計、例えば、1組の科学手法や広範な分野に適用できる方法にもとづいた大規模な流

動的なチーム、基礎科学から市場性のある技術にいたる研究を支援する管理インフラストラクチャを提示する。プロジェクトマネージャーは伝統的な研究代表者より産業研究開発マネージャーに類似していると思われる方法でこれらの多様な活動を掌握する。

基礎研究はしばしば応用的成果をもたらしてきた。基礎研究の順応性だけでなく、活動の他の側面に対する基礎研究の貢献についても明確な証拠が見出される。評価が研究開始の際に掲げられた目標という狭い範囲に集中する場合は、これらの広範な影響は見落とされることがある。

学際的な研究には小規模な、個々の研究者、専門分野指向のプロジェクトより多様なマネジメントが必要になってくる。BES のプロジェクトは一般的に多分野かつ研究の注目点を多数持つようなチームによって実施され、かつこれらのチームは伝統的なピアレビューに加えて、従来は見られないような組織や能力の形成などを重視する姿勢によって運営されている。

大学と連邦研究所のプロジェクトの結果は範囲、技術的な注目点および影響などが似ていた。大学は大学院生を利用できるが、仲間内の専門分野重視傾向が邪魔になっている。政府研究所はより安定性があり、より明瞭な中心をもっている。大学と連邦研究所の研究者はそれほど相互交流がなかった。

#### コメント

研究者はプロジェクトの定義に依存する評価方法の限界を強調している。プロジェクトが研究資金供与のために政府機関が行使する行政的便宜だとしても、それらは研究者が各自の研究に対して持っている見解とは一致しないことが見出された。研究者はもう少し長期的な視点を持ち、かつ長期的技術目標達成のため協調する人的資本やネットワークの形成に焦点を置くべきだと主張した。彼らの考えでは、こうしたもう少し長期的な方法をとらなければ、研究便益は過小評価を受けるだろうとしている。RVM(研究価値マッピング)プロジェクトが生産的なネットワークと非生産的なネットワークとのあいだの相違を明らかにすることができれば、彼らのネットワーク評価方法は科学研究の評価のため進歩した理論的に健全な基盤を作り上げることになる。

### 3. 技術分野の重要インフラストラクチャ特定のためのテキスト・マイニング

Ronald N. Kostoff,

“Implementation of Textual Data Mining in Government Organizations,” presented at:

Federal Data Mining Symposium and Exposition, Washington D.C. March 2000.

#### 背景

Ronald Kostoff は長年米国海軍研究局 (ONR) テクノロジーアセスメント局の局長を務め、自著の中で定量的評価手法を用いた。彼の最も新しい研究は、テキスト・マイニングの手法を科学技術のマネジメントに適用することにあつた。これらの手法はしばしば商業的かつ競争的な諜報業務で使用されているが、政府の評価研究の大半にはマイクロすぎて適さない。

#### 方法

Kostoff は 1998 年にテキスト・マイニング法をはじめて実施した。これには次のようにいくつかの段階がある。

- 1) 技術的テーマ(炭素分子体のフラレン、または船舶流体力学などの)関連の論文を特定するためのフィルター<sup>1</sup>の反復的開発。6つのデータベースがソースとして使用された。
- 2) 単語や句が文書に現われる頻度が割り出された。各項目の専門家が有効な句を選択した。
- 3) それぞれの有効な句について密接な関連をもつ句の辞書を作成し、これをもとに、その他すべての句が各有効句に接近して出てくる回数を集計した。各関連句には有効句への関連の強さの尺度が付与されている。しきい値を設定して最も密接に結びついた句を選び出す。各項目の専門家は辞書やそれらの概念関係の中からテーマを特定する。
- 4) データベース、テーマ間の関係および下位分野の関係を分析して、これらの中に頻繁に出てくる技術テーマを特定する。

---

<sup>1</sup> フィルターの検討については方法論の章を参照のこと

- 5) 著者、ジャーナル、所属などを使って計量文献学的分析を行い、テーマを研究者に関連付ける。

## 結果

計量文献学によって各技術分野の重要なインフラストラクチャの場所が特定できた。これはワークショップやレビューパネルのための専門家の発掘や、企画訪問をするのに役立った。さらに計量文献学によって生産性や影響を追跡することができ、かつ重要な知的遺産を特定することができる。基準がないため、専門分野をこえて計量文献学データを比較し、任意の分野における異常を探り当て、一般的傾向を特定するようにした。

専門家が句を技術的分類に収め、管理に役立てた。海軍要求文書の中の句を処理して上記の文献中の句と同じ分類法に収めた。要求文書、文献双方の分類区分に対する力点の置き方の定量的推定を行った。要求文書と文献の対照マトリックスによって専門家は各区分の技術的力点の妥当性や欠陥を判定することができた。各区分における技術的可能性レベルを推定する専門家に関しては同様な、しかしもう少し柔軟な方法がとられた。区分の分析手順は研究における重要なパラメータであった。もう少し細かい区分(例えば、「材料」よりむしろ「溶接チタン合金」)の方が有益だが、そうした作業は費用や時間がかかる。

当該方法を取ったのは主にコンピュータを使って作業をしている各項目の専門家だった。したがって、結論は専門家のバイアスと限界を反映している。データの最大異常件数を探知する信頼できる分析を行うためには多様な知識を備えた専門家が必要とされる。そしてジェネラリストは分野の専門家であれば特有と認めないかもしれない技術分野における特有なパターンを特定するのに必要である。組織の長期戦略的観点から、主たる成果は文書の作成ではなく、むしろ当該専門家の見解が広がることである。専門家はツールを使用して研究目的に取り組む方法や作成情報を分析し、かつ解釈する方法などを学ばなければならなかったが、学習カーブは勾配が急であった。

## コメント

Kostoff はすべての科学技術マネジメントの意思決定の手段は相互関係にあり、科学技術の戦略的マネジメントをサポートするために統合する必要があると考えている。したがって、プログラム・ピアレビューはマトリクスを使って戦略目標に対する進捗を測定すべきであり、ロードマップによりプログラムをもう少し大きい空間・時間の中に位置付けるべきであり、テキスト・マイニングによって確実にロードマップを優れたものにすべきである、などとしている。

#### 4. 米国研究分野のベンチマーキング

**Experiments in International Benchmarking of US Research Fields Committee on Science, Engineering, and Public Policy (COSEPUP), National Academy Press, Washington D.C., 2000**

Available at: [http://books.nap.edu/html/exp\\_in\\_bench/pdf/](http://books.nap.edu/html/exp_in_bench/pdf/)

##### 目的

米国の研究リーダーシップとしての地位を評価するためにベンチマーキング手法を用いる。国際ベンチマークにより世界基準を備えた1国(または1地域)における品質と影響を比較することができる。

##### 方法論

数学、免疫学、および材料科学・工学の3分野を選んだ。各分野では COSEPUP が報告書作成のため著名な科学者のパネルを任命した。パネルが用いた方法は次の通りであった。

- 仮想議会－パネルメンバーは下位分野の主要専門家に世界のベスト5から 20 に入る人は誰かをたずねた。その結果得られたリストには、指標を作成するために国別情報を加えた。
- 引用分析－各分野では既存の英国式分析が用いられた。免疫学パネルは ISI からハイインパクトな免疫学データベースを購入し、そして多産な著者を引用集計で格付けし、かつ国によって分類した。
- ジャーナル発表論文の分析－ジャーナル5誌が精査され、研究代表者の場所(所属)や各自の下位分野の一覧表が作成された。
- 定量的データ分析－一般に教育と資金供与に関するデータは国際的には比較できない。
- 受賞分析－各分野において主要な賞を得た米国人および非米国人の数を集計した。この場合受賞者の移動に留意した。
- 国際会議講演者－会議講演者のうちの米国人代表者を集計した。この場合、会議のオーガナイザーが地理的バランスに配慮して講演者を多様な国から招待することに留意した。

## 結果

各パネルは、米国が当該分野において世界のリーダーの一角を占めていると結論した。しかしながら、各パネルはまた米国が世界のリーダーに遅れをとっている下位分野を特定した。各パネルは問題となるインフラストラクチャを特定した。

## コメント

報告書には興味深い緊張が含まれている。例えば、報告書は定量的指標を非難している。このことは全米アカデミーの出版物において決まって問題にされる点である。報告書によれば、定量的指標は有益だが、それだけでは不十分であり、このため計量文献学に関する論拠を十分に再考しなければならないとしている。例えば、本当にイノベーティブな研究を記述する論文でも、他に匹敵する研究を行っているものが一人もいない場合、ほとんど引用されない。したがって、「パネルメンバーの専門的判断が最も効率的な研究評価をもたらした」(6-7頁)。しかし、必要であれば、パネルメンバーである科学者は計量文献学による情報を利用した。

報告書はまた費用効果を強調している。これは日常的に使用するために提案される方法論における重要な要素である。報告書作成には国の内外の非常に地位のある多くの科学者があたり、そして報告書の性格から、彼らはその執筆にアカデミー委員会業務で見られる以上の関与をしなければならなかった。しかし、彼らの費やした時間には報酬が支払われなかった。したがって、彼らの参加費用は彼らの所属機関や彼らに研究助成金を支給した政府機関がやむを得ず負担した(この方法が日常ベースで用いられるとすれば、外国政府が米国の科学政策策定作業に補助金を支給していることになる)。こうしたことから、その手続きには非常に費用がかかり、実費にはアカデミーへの直接請求だけを含めることによって過小請求された。この種の研究の総費用の計算については Kostoff のハンドブックを参照のこと。

各ベンチマークグループは米国が科学分野における世界的なリーダーであるかどうかを評価する作業を担当した。報告書は主観が不可避であることを認めつつも、それに取り組んでいる。非常に地位の高い米国人科学者に、彼らの仕事が当該分野での米国のリーダーシップに結び

ついたかどうかを判断するよう要請した。関与した科学者はむずかしい選択に迫られている。米国が世界のリーダーであると言え、多分、科学者にはもはやこれ以上資金は必要ないと断じる者がいるかもしれない。しかし、米国は世界のリーダーではないと言うのも間の悪い話である。とくに、米国は他のどの国よりはるかに大きな努力をしているからである。





CHIリサーチが米国連邦政府のために行った研究評価

**B.**

**論文間の分析: 研究結果 ( output ) の評価**

## 1. 研究目標の促進を求める法律の効果の評価

報告者: Samuel Reisher & Francis Narin, May 1980

### 目的

定量的指標の開発目的は、1972 年国家心臓・血管・肺・血液法の研究活動への効果を検証することだった。この 1972 年法は国立心臓・肺・血液研究所 (NHLBI) に心臓、肺、血管および血液などの病気と闘う広範な権限を与えた。重点的に分析を行ったのは、肺の構造と機能 (LSF)、慢性閉塞性肺疾患 (COPD)、および高血圧 (HPT) の3分野であった。

LSF と COPD はとくに 1972 年法の目標に設定され、他方 HPT はそれ以前の法律の目標に設定されていた。したがって、HPT は一種の対照分野の役割を果たし、COPD や LSF の研究の傾向はこれと比較して評価された。

### 方法論

当該研究に関する書誌は、3つのデータベースを使って作成された。これらデータベースの各々は論文に対する一意な情報を提供した。CHI は結局次の情報を含むデータベースを編纂した。すなわち、参考文献、機関、資金供与に対する謝辞、引用集計、研究レベル、ジャーナル影響度、主題分野である。上記研究3分野の論文を特定するために医学文献検索システム Medline の主題分類が使われた。顧客との共同作業で各研究分野の洗練された定義を開発し、約1万5,000 件の論文の 1 つひとつを分類した。各論文の主題が非常に詳細に検討されたので、論文の詳細な分類が開発され、内容分析が行われた。

### 結果

1972 年法の後、2つの目標分野における NHLBI のサポートを記載した論文の比率は上昇した。この効果は対照分野では見られなかった。さらに、2つの目標分野における NHLBI 論文の平均引用率が上昇し、それらが最も引用度の高い論文である可能性が拡大した。これも対照分野については見られなかった。

研究の結果、1972 年法は臨床医学研究への資金援助が減少する時期がに施行されたことがわかった。臨床医学論文の件数はこの頃多くの分野で減少し、目標分野(COPD および LSF)の規模は拡大しなかった。HPT は初期の法律の目標とされ、そして 1972 年法施行までは増加を続け、その後も引き続き増加した。

論文のその他の特徴は 1972 年法により影響を受けなかった。この間、検証された論文はもう少し基礎的な臨床医学研究にシフトしたが、目標分野および対照分野の論文はともに同様な傾向を示した。内容分析により、(目標分野および対照分野いずれにおいても)NHLBI の資金を供与された研究が先導する生理活性物質の同定や評価へ力点を強化していることがわかった。これは NIH が一般的に先導している臨床医学研究のもう少し科学的なタイプへの重要なシフトであった。

#### コメント

ここでは研究関連法規の効果を明らかにするために計量文献学を利用した。非常に微妙な効果が明らかにされた。

## 2. 研究所の出版物の包括的なベンチマーキング

報告者: Paul McAllister & Francis Narin, October 1982

### 目的

公開発表文献を利用して海軍研究所の科学的成果の定量的指標を開発する。

### 方法論

米国軍の個別の研究技術開発、工学施設をはじめとする 500 の個別論文発表ユニットが特定された。さらに、米国非軍事連邦機関 19、FFRDC (Federally Funded Research and Development Centers、連邦出資研究開発センター) 18 および企業 30 社が個別に特定され、分析された。データには米国の非大学論文の主要部分のほとんどすべてが含まれていた。各論文発表ユニットの研究の発表規模、主題設定および引用の成績などの各論文発表ユニットの3部からなる評価が実施された。これらの評価には工夫を施し、非常に異なった調査分野に従事する部門全体にわたって完全に比較できるようにした。

### 結果

国防省 (DOD) の全施設の中での海軍研究所 (NRL) がトップであることは明白だった。NRL は DOD の中で単独で最大の論文発表ユニットで、オークリッジまたはアーゴンなどの主要連邦研究所に匹敵していた。NRL の科学者は物理学、地球および宇宙科学・工学で活躍していた。NRL の著作はまた非常に引用度が高く、すべての海軍施設の中で最高の引用を誇っていた。とくに海洋学は顕著だった。NRL の引用得点が劣っている下位分野もいくつか確認された。海軍の論文は空軍より引用度が高く、そして NRL はこれに大きく貢献していた。2年間にわたり、9 つの主要分野の各々において引用度が最も高かった上位 25 の論文が決定され (計 450 点)、そして海軍の著者はこれらのうち 18 件に関わっていた (表 4-1 を参照のこと)。DOD の科学者は一般に期待されるほど引用度は高くなかったが、その理由は彼らの研究の専門的性格によるものだった。

## コメント

本研究は計量文献学を適用して論文発表を主たる任務としない研究所の成果の出版の要素を検証した。本研究はベンチマーキング作業に機関を持ち込んだという点で優れている。

表 4-1 Navy Papers in the Top 25 Most Highly Cited in Their Field

<u>Field</u>	<u>Rank in the Field</u>	<u>Number of Citations</u>	<u>Journal</u>	<u>Authors</u>	<u>Title</u>	<u>Corporate Address</u>
Bioned Res	21	82	Science	Nelsonre. WA Flanders. RR Hawthorn. PK	Banded Marker Chromosoms as Indicators of Intra- species Cellular Contamination	USN, Biomed Res Lab Univ Calif, Sch Publ Hlth, Cell Cu
Earth & Space	12	50	J Geoph Res	Vogt. PR Avery. OE	Detailed Magnetic Surveys in Northeast Atlantic and Labrador Sea	USN, Oceanogr Off
Eng & Tech	5	35	P IEEE	Taylor. HF Yariv. A	Guided Wave Optics	USN, Electr Lab Ctr Caltech, Dept Elect Engn
	7	31	Computer Ph	Manning. I Mueller. GP	Depth Distribution of Energy Deposition by Ion- Bombardment	USN, Nucl Sci Div, Res Lab
	22	22	IEEE MICR T	Granatst. VL Herndon. M Parker. RK Schlesin. SP	Strong Submillimeter Radiation from Intense Relativistic Electron Beams	USN, Res Lab
Psychology	8	16	Psychophysl	Johnson. LC Naitoh. P Moses. JM Lubin. A	Interaction of REM Deprivation with Total Sleep Loss- Experiment 2	USN Hosp Neuropsychiat Res Unit
	9	15	Psychophysl	Lubin. A Moses. JM Johnson. LC Naitoh. P	Recuperative Effects of REM-Sleep and Stage 4 Sleep on Human Performance After Complete Sleep Loss- Experiment 1	USN Med Neuropsychiat Res Unit
	23	11	Pharm Bio B	Shearer. DE Fleming. DE Bigler. DE Wilson. CD	Suppression of Photically Evoked After Discharge Bursting Following Administration of Anti- convulsants in Waking Rats	USN, Undersea Ctr Vet Admin Hosp, Neuropsychol Res Lab Brigham Young Univ, Dept Psychol
Mathematics	9	6	Am Math Mo	Gaskell. RE Klamkin. MS	Industrial Mathematician Views His Profession- Report of Committee on Corporate Numbers	Ford Motor Co., Scientific Res Sta USN, Post grad Sch, Dept Math
	16	5	Math Comput	Neild. C	3--Rank of Quadratic Fields and Euler Product	USN, Ship Res & Dev Ctr, Comp & Ma

### 3. 資金供与メカニズムの比較:センター・プログラム方式とピアレビュー方式による資金供与

報告者: Samuel R. Reisher & Francis Narin, July 1984

#### 目的

本研究はセンター・プログラム方式による資金供与を評価し、このためセンターの科学的結果(output)を伝統的な研究提案公募による研究助成金の資金供与を受けた科学者の成果と比較した。本研究が評価しようとしたのは、センター・プログラムがその当所の目標を達成していたかどうか、およびピアレビューによる資金供与からの結果(output)と比べて科学的品質に何らかの見劣りがあったかどうかという点だった。

#### 方法論

センター、センターの研究代表者および R01 グラントの受領者が発表した論文のデータベースは、顧客が提供した情報を2つの論文データベースと結びつけて苦労してまとめあげたものであった。CHI はこの引用データに SCI(科学引用索引)から抽出した引用機関データや下位分野の分類、基礎から応用までの研究レベルの分類や資金援助への謝辞などを加えた。センター論文、センターの研究代表者による論文および R01 受領者による論文などをさまざまに組合せて7つの相互に独立した論文区分を作成した。これらの区分の計量文献学的特性がセンターの目標に関連して分析された。

#### 結果

センターは、基礎的臨床医学研究の研究背景や関心をもつ広範な科学者を引き付けたことがわかった。センターの科学者は NIH の他の部門からの支援を含めた追加資金供与源からも資金を得た。科学者はセンターに参加後、引き続き広範な下位分野の論文を発行している。センターは個人間の協力拡大を促進した。センター科学者の論文は、当該センター方式以外の機関の論文から引用されている。センター科学者は多作の人が多く、また彼らの論文は R01 の研究代表者の論文より引用度が高い。目標を計量文献学特性と評価結果に関連付けた表 4-2 を参照のこと。



## コメント

本研究はプログラムがその事前設定目標を達成しているかどうかを定量的に証明するために計量文献学を使う好例である。本研究は多様な目標をこの手法で評価できることを示している。

## 訳注)

NIH のグラントには多くの種類があり、1つのアルファベットと2桁の数字で表されている。

P で始まるシリーズが Research Program and Centers と呼ばれるセンター方式である。この中で最も一般的な P01 (Research Program Project) は、特定の目的やテーマのために異分野の研究者を集めてセンターを作り、長期に渡ってマルチディシプリナリーに研究を進めるものである。

個人やグループで研究課題申請を行う一般的なグラントが R シリーズ (Research Projects) であり、NIH グラントの多くは R01 (Research Project Grants) として支給されている。研究者による申請課題に対して、その所属する機関(国内外問わず)に資金供与を行う。

表 4-2 – Policy goals, bibliometric measures & results in Center-R01 study

Center program goal	Bibliometric measures	Results
	research level	<ul style="list-style-type: none"> <li>Center papers were much more oriented to basic research than R01 supported papers.</li> </ul>
To attract basic biomedical scientists to develop a more interdisciplinary approach to center research.	subfields	<ul style="list-style-type: none"> <li>Center: 30% of papers appear in biochemistry &amp; molecular biology subfield and 20% in center focus. R01: 50–60% of papers in center focus, 10% in biochemistry &amp; molecular biology.</li> <li>New principal investigators (PI's) in centers more often published primarily in fields other than center focus as compared to new R01 P.I.'s</li> </ul>
	citing papers	<ul style="list-style-type: none"> <li>R01 papers were more likely to be cited by center focus papers and by papers acknowledging NIDR support. Center papers were more likely to be cited by papers produced at other centers.</li> </ul>
Attract new types of money to dental re-search	support acknowledgements	<ul style="list-style-type: none"> <li>Center papers acknowledged about twice as many sources of support as R01 papers</li> </ul>
Attract new people to dental research	individual's publishing histories	<ul style="list-style-type: none"> <li>Centers recruited more scientists who published in the basic bio-medical sciences.</li> <li>Centers &amp; R01 were equally effective in recruiting new scientists, those without publications in the previous 6 years.</li> </ul>
Encourage collaboration	number of co-authors	<ul style="list-style-type: none"> <li>Center papers averaged 3.1 authors; R01 2.9 authors. Papers with 4 or more authors comprised 33% of center papers and 26% of R01 pa-pers.</li> </ul>
	collaborating institutions	<ul style="list-style-type: none"> <li>Both averaged 1.3 institutions per paper</li> </ul>
Produce research of as high a quality as produced when each project is individually peer-reviewed (R01 mechanism)	number of papers	<ul style="list-style-type: none"> <li>Center and R01 scientists similar. Those who have both center and R01 support publish about 50% more papers per year than those with either center or R01 funding.</li> </ul>
	number of citations (controlled for year and subfield)	<ul style="list-style-type: none"> <li>Center and R01 scientists similar. Those who have both center and R01 support are more highly cited than those with either center or R01 funding.</li> </ul>

#### 4 研究機関における日常業務としての計量文献学的プロファイリングの例

報告者: Francis Narin および CHI スタッフ (作成日の記載なし)

##### 目的

国立児童保健・人的開発研究所 (NICHD) の研究助成による論文の分析的評価。同様の報告書が NIH の各部門のために作成された。

##### 方法論

進行中の作業の一環として、本報告書は CHI が NIH のために作成した計量文献学の専門データベースを活用した。このデータベースには主要臨床医療ジャーナルにおける論文の資金援助への謝辞に関する情報が含まれていた。これは SCI から得た引用や機関情報と組み合わせられた。これらの報告書について CHI は各論文に関する次の情報を含む分析的データベースを編纂した。参照文献、機関、資金援助への謝辞、引用集計、研究レベル、ジャーナル影響度、主題分野。さらに、相互引用分析のため引用の連関が利用された。

##### 結果

NICHD はある期間にわたって臨床的研究と基礎的研究をかなりコンスタントな割合で組合せた活動を支援した。NICHD 論文の件数は約3%減少した。当該研究所の主たる影響は内分泌学、産科学、婦人科学、小児科学、解剖学および形態学並びに発生学などの下位分野にあった。NICHD は 1973 年から 80 年の期間、発生学における論文の 12%、および解剖学と形態学における論文の 10%を占めた。生化学や分子生物学への支援は強化されたが (NICHD が支援した論文の 17%)、同研究所は当該下位分野における論文の 1.6%を占めるに過ぎない。

NICHD の「活動度指標」、言い換えれば当該下位分野におけるすべての論文に対する同研究所の下位分野における論文の尺度は、発生学関連が最も高く、1977 年～80 年には 7.2、すなわち、「見込み」レベルの 7.2 倍であった。同研究所は次の下位分野においてはその他一切の NIH

研究所より多くの論文を支援した。受胎能力、産婦人科学、小児科学、解剖・形態学および発生学。

NICHD 支援文献における大変更を示す下位分野は次の通りであった。栄養学および食餌療法学－33%増、細胞生物学/細胞学/組織学－31%増、および受胎能力－19%減。老年医学の研究は1974年における全米加齢問題研究所の創設とともに衰退した。

NICHD の顕著な研究者への支援は同研究所の論文への引用を見れば明らかである。同研究所が最も活動している16の下位分野のうち15分野においては、その論文の10%以上は被引用論文の第10十分位にあった。これは同研究所のプログラムの品質が当時資金援助制限をしていたにもかかわらず上昇し、維持されたことを示している。

NIH 研究所群が各々の研究において相互に依存していることは引用が相互に交流していることから明らかである。NICHD が支援している論文の中で当該研究所が支援している研究に引用されるのはその論文における参考文献の20%にすぎない。残りは他の機関が支援している論文に対する引用であった。逆にNICHD の支援する研究への引用の84%は他の支援による論文に掲載された。NICHD、NIADDK、NCI および NIGMS の間の相互依存関係は強かった。

いくつかの図と表を例示のため掲載してある。図4-1 および 4-2 を参照のこと。また同研究所の計量文献学活動を要約した論文発表のプロファイルについては図4-3を参照のこと。

## コメント

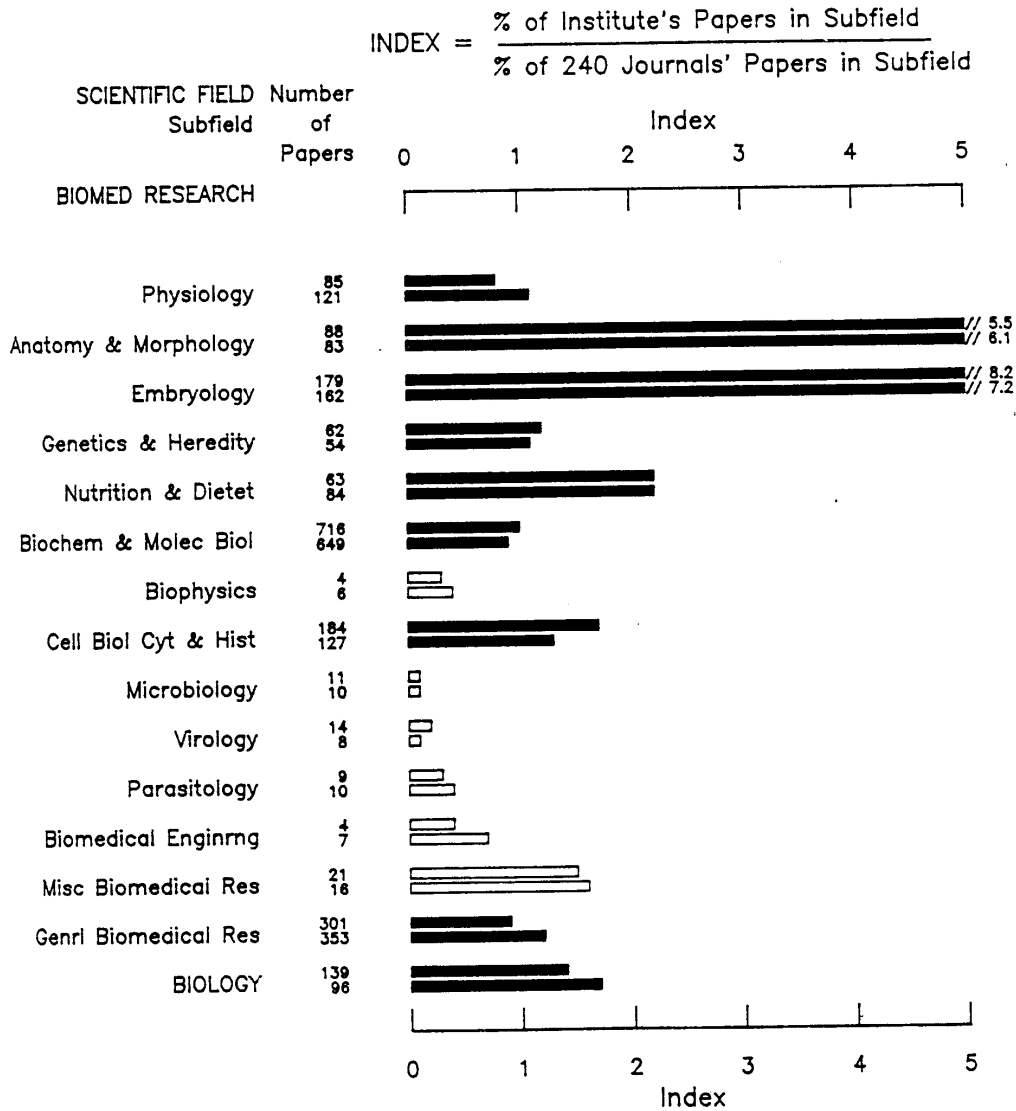
本報告書は比較可能な機関プロファイルを提供するために継続的に利用されている計量文献学の1例である。NIH の資金供与が比較的安定しており、かつそれらの機関が最高レベルの科学的優秀さで運営されていることは多分偶然ではないだろう。したがって、NIH は定量的尺度を通して短所が露見することをほとんど恐れていない。

☒ 4-1

NICHD Activity Indexes in Subfields of Biomedical Research

240 BID-MEDLINE Journals

1973-1976 over 1977-1980



Note: Open bars designate subfields that contained less than 1 percent of the papers acknowledging this Institute's support in the

4-2

Universe of Publications Cited by and Citing NICHD-Supported Papers

Area of Circles Proportional to Number of Publications Supported; Citations by Percent

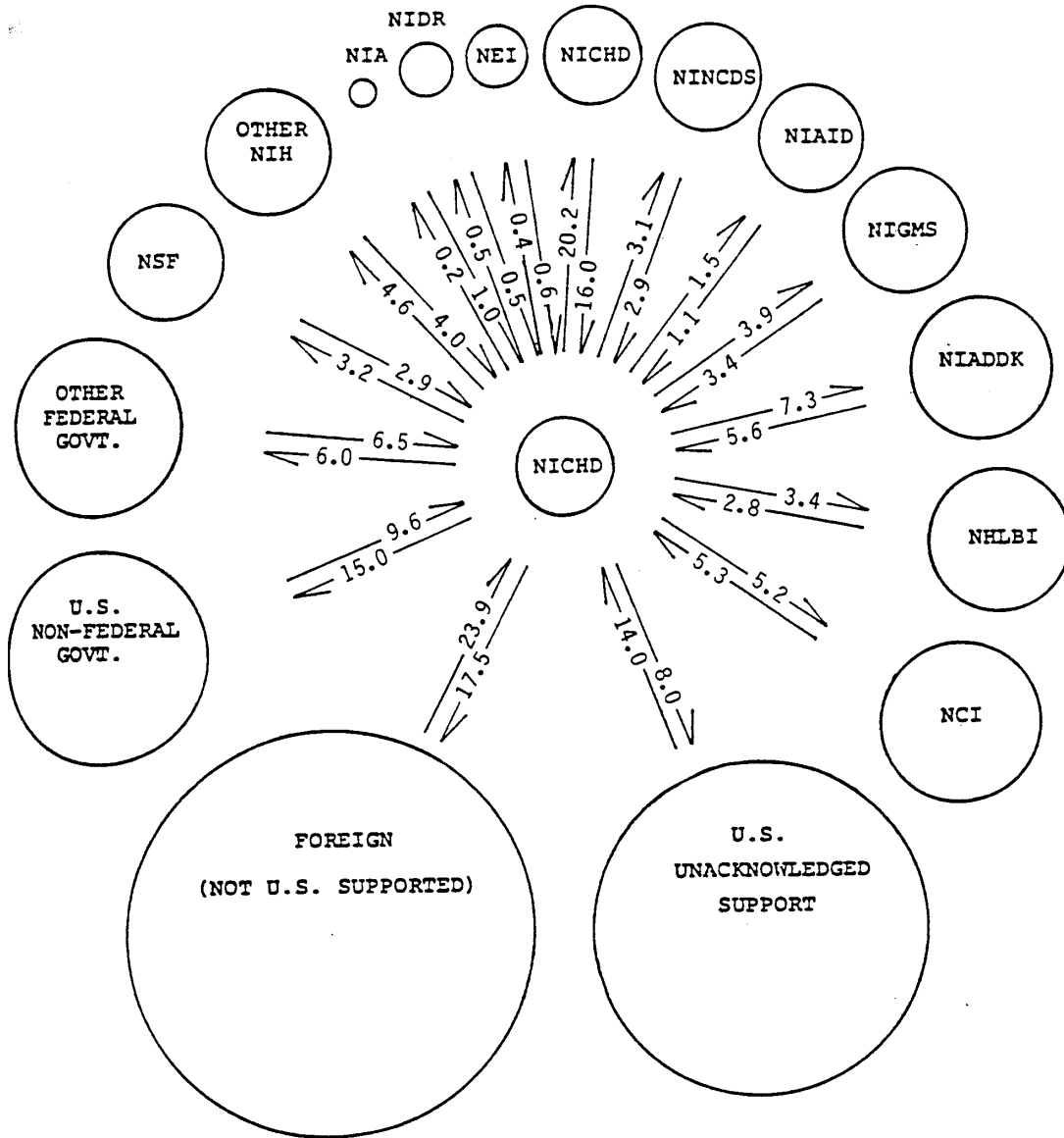


表 4-3

Bibliometric Profile of NICHD for the Cited Years 1977-80

Citation Counts Scaled to 1973 Equivalents

Based on Cites Received from

All Citing Years, 1973-80\*

FIELD/LEVEL	NUMBER PAPERS	ACTIV INDEX	% INT EFFRT	% EXT EFFRT	TOTAL CITES	CITES PER PAPER	CITE T-SCR	% PAPERS AMONG TOP 10%
***** NORMALIZED AGAINST FIELDS [1-2] *****								
FIELDS [1-2]	3892.	1.00	100.00	1.70	23929.	12.0	52.01	12.41
CLIN MEDICINE [1]	2203.	0.93	56.60	1.57	11387.	10.3	53.63	16.25
BIOMED RESEARCH [2]	1689.	1.12	43.40	1.90	12477.	14.0	50.46	10.75
LEVEL								
CLINICAL-MIX [1-2]	773.	0.59	19.85	1.00	3512.	8.5	53.71	18.96
CLIN INVESTIGATN [3]	1336.	1.39	34.32	2.35	7098.	11.0	52.54	12.79
BASIC RESEARCH [4]	1784.	1.10	45.83	1.87	13196.	14.1	50.42	10.71
***** NORMALIZED AGAINST FIELDS [1-9] *****								
FIELDS [1-9]	4226.	1.00	100.00	1.77	24407.	11.2	51.68	12.68
CLIN MEDICINE [1]	2203.	0.89	52.13	1.57	11387.	10.3	53.63	16.25
BIOMED RESEARCH [2]	1689.	1.07	39.97	1.90	12477.	14.0	50.46	10.75
BIOLOGY [3]	96.	1.73	2.26	3.06	310.	5.7	54.62	22.87
CHEMISTRY [4]	18.	0.36	0.42	0.64	87.	9.1	52.34	3.51
PHYSICS [5]	13.	1.78	0.31	3.15	40.	6.3	53.58	21.05
EARTH & SPACE SC [6]	0.	0.00	0.00	0.00	0.	0.0	0.00	0.00
ENGRNG & TECHNOL [7]	0.	0.00	0.00	0.00	0.	0.0	0.00	0.00
PSYCHOLOGY [8]	208.	3.29	4.91	5.81	492.	4.2	54.26	18.31
MATHEMATICS [9]	0.	0.00	0.00	0.00	0.	0.0	0.00	0.00

NOTE: CHEMISTRY AND PSYCHOLOGY DATA SHOULD BE TREATED CAUTIOUSLY OVER TIME DUE TO INCONSISTENT DATABASE COVERAGE.

\*1979-80 papers not included in present citation data.

表 4-3 (続き)

Bibliometric Profile of NICHD for the Cited Years 1977-80

Citation Counts Scaled to 1973 Equivalents

Based on Cites Received from

All Citing Years, 1973-80\*

(Subfields Normalized Against Fields 1-2)

FIELD/SUBFIELD	NUMBER PAPERS	ACTIV INDEX	% INT EFFRT	% EXT EFFRT	TOTAL CITES	CITES PER PAPER	CITE T-SCR	% PAPERS AMONG TOP 10%
<b>CLIN MEDICINE [1]</b>								
GENRL & INTERNAL MED	221.	0.68	5.67	1.15	1595.	13.4	53.29	12.78
ALLERGY	3.	0.31	0.06	0.53	2.	2.6	38.12	0.00
ANESTHESIOLOGY	7.	0.31	0.18	0.53	14.	2.8	47.01	20.00
CANCER	37.	0.23	0.95	0.39	304.	13.6	52.54	9.70
CARDIOVASCULAR SYSTM	8.	0.08	0.20	0.13	33.	6.9	43.61	0.00
DENTISTRY	4.	0.07	0.11	0.11	1.	0.4	44.20	0.00
DERMAT & VENERL DIS	7.	0.24	0.17	0.40	22.	5.8	55.89	21.74
ENDOCRINOLOGY	622.	5.48	15.99	9.29	4284.	15.5	58.57	14.94
FERTILITY	160.	5.15	4.12	8.73	483.	5.7	52.01	18.58
GASTROENTEROLOGY	7.	0.20	0.17	0.33	23.	8.8	54.63	12.50
GERIATRICS	12.	0.65	0.32	1.10	20.	2.7	54.49	25.58
HEMATOLOGY	7.	0.16	0.18	0.27	12.	5.6	45.31	0.00
IMMUNOLOGY	91.	0.51	2.35	0.87	572.	10.8	46.24	6.58
OBSTETRICS & GYNECOL	206.	3.57	5.29	6.06	837.	8.3	55.27	24.09
NEUROL & NEUROSURG	241.	0.98	6.19	1.67	1248.	10.2	50.17	9.44
OPHTHALMOLOGY	9.	0.17	0.22	0.28	72.	14.0	68.94	35.48
ORTHOPEDICS	3.	0.12	0.08	0.20	0.	0.0	44.64	0.00
ARTHRITIS & RHEUMAT	3.	0.18	0.08	0.30	3.	2.1	42.71	0.00
OTORHINOLARYNGOLOGY	17.	0.51	0.45	0.87	61.	4.2	54.12	24.14
PATHOLOGY	20.	0.39	0.51	0.66	69.	7.4	54.11	25.00
PEDIATRICS	315.	3.52	8.09	5.97	1148.	7.4	54.60	19.89
PHARMACOLOGY	77.	0.49	1.97	0.82	365.	9.6	50.91	14.10
PHARMACY	25.	0.47	0.64	0.80	102.	6.4	53.92	20.83
PSYCHIATRY	14.	0.24	0.36	0.41	22.	3.1	46.13	9.30
RADIOLOGY & NUCL MED	8.	0.08	0.20	0.13	14.	4.3	49.53	15.00
RESPIRATORY SYSTEM	10.	0.47	0.27	0.80	60.	9.2	67.48	30.77
SURGERY	8.	0.05	0.20	0.08	8.	2.6	45.40	0.00
TROPICAL MEDICINE	2.	0.13	0.04	0.21	0.	0.0	0.00	0.00
UROLOGY	13.	0.30	0.32	0.51	35.	4.2	52.51	18.00
NEPHROLOGY	0.	0.00	0.00	0.00	0.	0.0	0.00	0.00
VETERINARY MEDICINE	16.	0.33	0.40	0.56	55.	6.5	63.36	35.29
ADDICTIVE DISEASES	0.	0.00	0.00	0.00	0.	0.0	0.00	0.00
HYGIENE & PUBL HLTH	32.	1.21	0.83	2.05	76.	4.1	53.72	19.09
<b>BIOMED RESEARCH [2]</b>								
PHYSIOLOGY	121.	1.11	3.11	1.88	833.	14.7	51.35	9.97
ANATOMY & MORPHOLOGY	83.	6.06	2.13	10.28	320.	6.9	54.27	19.49
EMBRYOLOGY	162.	7.21	4.17	12.23	1004.	12.0	57.87	16.73
GENETICS & HEREDITY	54.	1.14	1.38	1.94	188.	6.4	45.57	7.95
NUTRITION & DIETET	84.	2.20	2.16	3.73	302.	6.8	50.48	8.24
BIOCHEM & MOLEC BIOL	649.	0.94	16.67	1.60	5419.	15.4	50.43	12.84
BIOPHYSICS	6.	0.42	0.16	0.71	18.	5.9	48.87	0.00
CELL BIOL CYT & HIST	127.	1.25	3.25	2.12	1047.	14.9	53.23	10.21
MICROBIOLOGY	10.	0.13	0.25	0.23	49.	10.9	57.64	29.63
VIROLOGY	8.	0.11	0.21	0.19	83.	20.7	65.70	25.00
PARASITOLOGY	10.	0.41	0.25	0.70	35.	6.7	61.31	25.81
BIOMEDICAL ENGINRNG	7.	0.67	0.18	1.14	12.	2.9	54.28	24.00
MISC BIOMEDICAL RES	16.	1.56	0.42	2.64	18.	1.8	49.99	8.20
GENRL BIOMEDICAL RES	353.	1.22	9.06	2.08	3280.	18.2	51.59	12.36



## 5. 重要ながん研究上の発見に影響する要因の評価

報告者: Francis Narin, September 1987

### 目的

国立がん研究所(NCI)は、がんの原因、予防、診断および治療に関連した研究を実施し、かつ支援している。本研究の目的は以下の点を確定し、これによって研究活動、資金供与メカニズムおよび実施者の場所などのあいだに重要な関係があるかどうかを確定することである。(1) 1965年～1982年の期間において選定されたがん研究の進歩に主に貢献した者は誰か？(2) 彼らはどのように支援を受けたか(契約、助成金、センター助成のいずれか)？(3) どこで研究を行ったか？

### 方法論

専門家パネルががん研究における 13 の重要進歩を選定し、そして各進歩の基礎になっている主要な研究活動を特定した。研究活動は、年代別およびテーマ別に配列し、各活動の歴史的トレースができるようにした(図 4-3 参照のこと)。各活動は1組の主要文書(その 90%は論文)に関連付けられ、がん研究における重要な進歩を示した。専門パネルは、改訂、完成されたトレース分析結果を評価し、コメントした。次の 4 組の論文群に整理した。1) トレースすべき論文そのもの、2) 密接な関連をもつ同時期の論文(共引用論文)、3) トレースしている論文によって引用された先駆的な論文、4) 関連性が低い、同時期の論文。各論文は次の事項で特徴づけた。支援源(NCI、他の NIH、米国の他の機関など)、メカニズム(助成金、契約、センター助成など)、研究場所(ハーバード大学、NCI 内部部門など)および引用回数。

### 結果

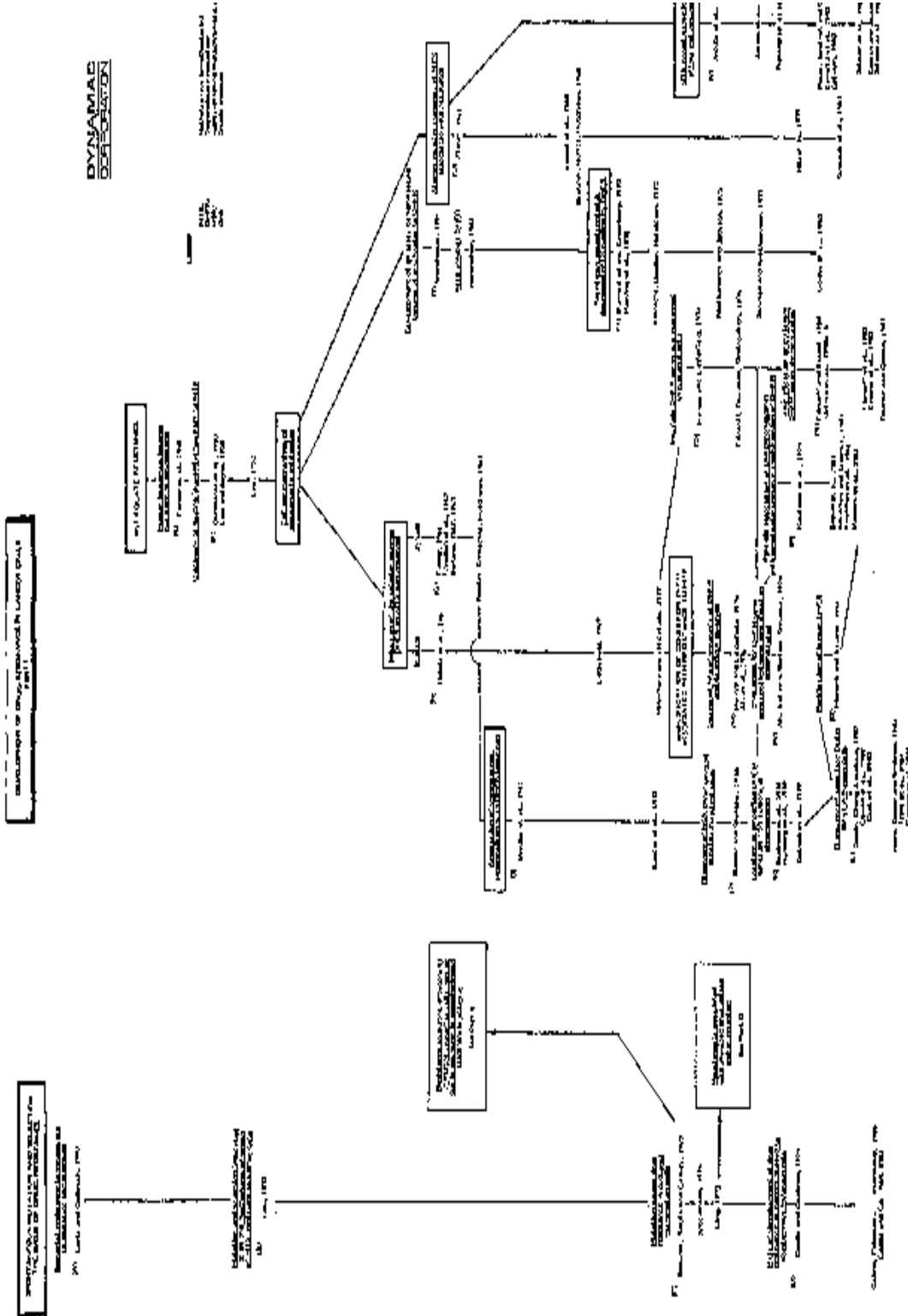
NCI 支援論文は、より広範ながん研究を集める周辺論文よりトレース論文に多く集まった。NCI が用いた主要な支援メカニズムの各々が同じ尺度で重要ながん研究結果やその周辺論文に貢献した。すべてのトレース論文はその下位分野や出版年の中で比較して非常に引用度が高かつ

た。重要な研究進歩をもたらしたのは NCI 自体はもとより大小の研究機関、米国および海外の機関、大学および医学部だった。こうして、支援のメカニズムや機関的背景は、研究の進歩を左右する主要な要因ではなく、そして言うまでもなく、NCI における研究申請の選択に用いた手続きにより、支援のメカニズムや機関の背景がどうであれ、影響度の高い研究が選定されつつあった。

#### コメント

本研究はトレース法と計量文献学のユニークな組合せである。もし、もう少し規格化されていたら、生じた結果はもっと確実なものになっただろう。すなわち、さまざまなメカニズムに対する NCI 資金供与の比率は、各タイプの資金供与を報告する論文の比率と比較することができ、かつ機関の論文発表規模はトレースされた論文の点数と比較することが出来たであろう。

TRACE #2: Development of Drug Resistance in Cancer



## 6. 内部的または外部的な研究支援メカニズムの比較

報告者: James Corrigan & Francis Narin, October 1989

### 目的

NIH の所内研究と所外研究の特徴の比較

### 方法論

進行中の作業の一環として、本報告書は CHI が NIH のために作成した計量文献学の専門データベースを活用した。このデータベースには主要臨床医療ジャーナルに掲載された論文における資金供与に関する情報が含まれていた。これは SCI から得た引用や機関情報と組み合わせられた。これらの報告書について CHI は各論文に関して次の情報を含む分析的データベースを編纂した。参照文献、機関、資金供与への謝辞、被引用数、研究レベル、ジャーナル影響度、主題分野。

### 結果

NIH 関連論文において、やや科学的な医療研究にシフトしたことは臨床医学を犠牲にして生物医学のより基礎的な下位分野を重視する傾向が強まってきたことから明らかになった。臨床医学分野における論文は、生物医学研究より引用パフォーマンスが相対的に高いことを示す傾向があった。

所外論文はデータベースの比率としては 15%で、所内論文は 1.5%であった。所外と所内双方の 1論文当たりの平均被引用回数は、データベース平均を十分に上回っており、全論文のそれぞれ 1.7 倍および 2.5 倍の引用を受けている。所外論文の 20%および所内論文の 30%は、それらの下位分野における上位 10%の最高引用文献の中に含まれていた(詳細は付属の図4-4 を参照のこと)。さまざまな引用実績尺度にもとづく一貫した格付けの順序付けは次の通りであった。NIH 所内論文 > NIH 所外論文 > 米国の全論文 > データベース中の全論文。

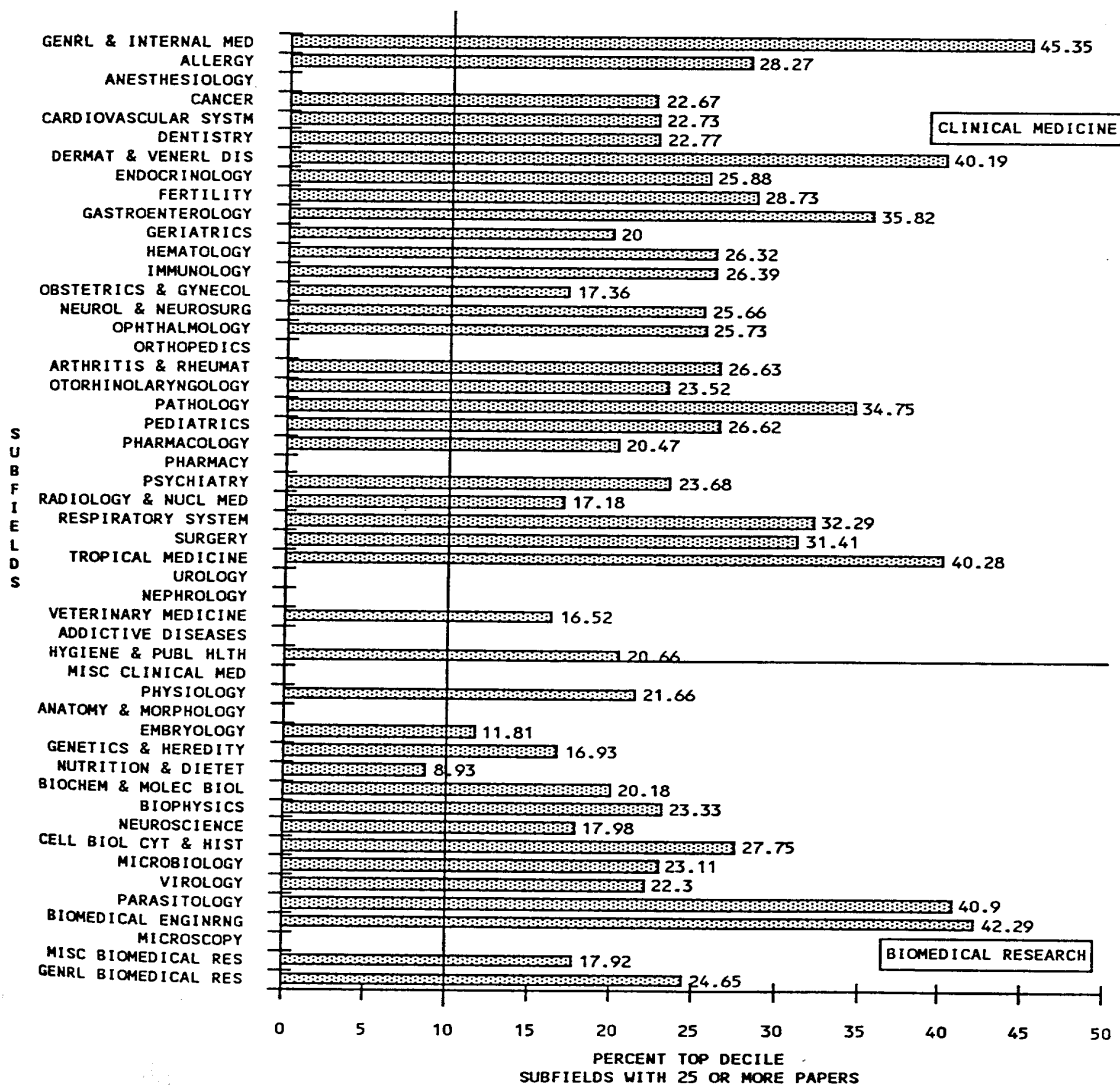
コメント

本項では2つの支援メカニズムが評価された。所内は NIH 自身の研究所を意味し、所外はピアレビューによる助成金を通して大学の研究者に供与された NIH 資金を意味する。これは進行中の作業の一部であった。そしてこの作業期間中に専門データベースが作成され、次いで特定の問題の検証や NIH 各研究所での定例報告書の作成などに利用された。

図 4-4

All NIH Intramural Percent of Papers in Top Decile by Subfield Publication Years: 1981-1984

Comboned Journal Set: Fixed 1981 NIH-Medline



## 7. 助成金受領者の成果と助成金申請が不採択だった者の成果の比較

報告者: James G. Corrigan & Francis Narin, 1990

### 目的

次の3組の論文の引用度を比較した。政府機関助成金受領者が作成した論文、政府機関によりプロポーザルが不採択となった者が発表した論文、および対照論文群。政府機関が資金を供与した研究が供与されなかった研究より引用度が高かったかどうか論じられた。

### 方法論

政府機関はプロポーザルリスト、研究代表者の名前およびプロポーザルと関連付けられた書誌などを提供した。CHI はこれらの資料を使って著者リストを作成した(様式の変更や書誌に関する情報の不備のため、これらの資料は著者やジャーナルの名称の収集に使用されたが、文献参照のソースとしては使用されなかった)。CHI はこれらの著者が出版した論文、ノートおよびレビュー論文など、言い換えれば申請者個人の書誌データの編纂上の通常の事項を探した。CHI は図書館で各論文を調べて資金供与に対する謝辞などを収集した。政府機関の支援に謝辞を呈する論文だけが集められた。これらの各々について、政府機関の資金供与を受けず、かつ同じ主題に関する対照論文が同様な量で、同じジャーナルから選択された(ジャーナルが学際的な性格のものである場合)。引用データも収集された。

### 結果

機関の5部門と出版年3年を検証し、合計 15 部門・出版年の組合せとなった。事例 15 件のうち 12 件において助成金取得者が発表した論文の中位被引用数は資金供与が不採択となった論文の中位数と対照論文の中位被引用数いずれよりも大きいかこれに等しかった。「完全な」記録を持っている者、すなわち、3年すべてにわたって資金供与が不採択となった者や対象者を上回る中位被引用数を持っている資金供与受領者がいるという点で部門には相違がでてきた。1つの部門

の資金供与受領者は非受領者より中位被引用数が少なかった。こうした部門は別の資金供与機関が有力な分野の論文に資金を供与した。

#### コメント

本研究は機関による資金供与研究との比較対象として同等の科学論文の設定を比較的大胆に行った。資金供与を拒否された著作を検証することは異例のことで、前述のように貴重な情報を与えてくれた。

## 8. 技術移転プログラムが科学者の論文発表に及ぼす影響の評価

Francis Narin & Kimberly Stevens Hamilton, April 1994

### 目的

CRADA、すなわち、共同研究開発協定 (Cooperative Research and Development Agreement) は会社と政府研究所とのあいだに結ばれる、共同プロジェクトの実施のための書面による協定である。CRADA は連邦研究所において開発された研究と技術の商業化を加速するために連邦議会が案出した手段である。本研究の根底にあるのは、CRADA を締結する科学者の質や当該協定が科学的成果 (output) を左右するかどうか、などについての問題であった。

### 方法論

本研究は CRADA を結んでいる NIH の科学者と結んでいない科学者の論文発表記録を比較した。さらに CRADA 締結前後の論文発表率も分析して、成果 (output) が低下している証拠を求めた。CRADA を結んでいる科学者の各々について、顧客は同じかできるだけ近い機関や格付けから CRADA を結んでいない科学者を選んだ。これらの照合された科学者は対照群を形成した。CHI は Medline の CD-ROM バージョンを使用して、116 人の科学者で構成する2つのグループの論文発表記録を編纂した。この過程は長く、問題含みだった。同名異人物は、データベースに姓や頭文字だけ記載されているために生ずる問題である。さらに、Medline は 1 文献当たり1つの住所だけを記載するため問題の解決にはどうしても住所情報は利用できなかった。文献の主題はそれが当の科学者に属するものかどうかを決定するのに使用された。当該研究のために書誌情報を編纂する上でこうした何万という判断が必要であり、そしてその属性は完全ではあり得なかった。

### 結果

本研究の結果わかったことは、CRADA を結ぶ科学者は対照科学者の2倍の論文を発表することだった。CRADA 科学者は年間8点から 12 点の論文を共著で発表する一方、対照科学者は4



点から6点だった。科学者が CRADA を結んだ後論文発表率がやや低下するとは結論できないが、そのいくつかの証拠はあった。対照科学者のグループは、統計的に有為ではないが、概して論文発表率の低下の様子は見せなかった(両側検定で 0.2 から 0.25 程度の有意水準)。

#### コメント

本項で説明した個別書誌の作成を考慮すると、結論を保証するための非常に強力な結果、例えば、CRADA 科学者が非常に生産性が高いというような結果が必要になる。CRADA 締結後の論文発表率の追跡のむずかしさに加えて本研究のタイミングの問題があった。最適の事例では、最初の CRADA の7年後に研究を行うとしても、CRADA 締結後に論文発表するまでに3年の猶予しかなかった(複数年のプログラム運営を検証したので、論文発表のずれを考慮に入れなければならない、かつジャーナル出版の遅れのため現行年の出版は利用できない)。このことは評価に特有の緊張関係、すなわち、意思決定者は早いフィードバックを好むが、研究がその分野に影響するには多年を要し、かつ技術に影響するには数十年かかるという緊張関係を示している。論文発表のずれは研究が影響するのに必要な時間を示しているが、これは計量文献学の短所と見られる場合が多い。

## 9. 世界的な競争者に対する研究所の発表論文のベンチマーキング

報告者: Anthony F. Breitzman, September 1995

### 目的

本研究は 1985 年から 1994 年までのあいだに発表された研究所の論文を評価した。そしてこのため多様なパフォーマンス指標を使って、当該研究所の長所と短所を決定し、当該研究所のパフォーマンスを他の同様な機関と比較した。

### 方法論

研究所の住所のすべての表現形を含む書誌データを SCI から編纂した。次いで、当該研究所の代表がどの論文が研究所から発表されたものかを決定し、当該論文を研究所部門別に分類するのを手伝った。非参照論文(会議抄録、議事録、年表など)は取り除かれた。

研究所各部門について次の指標を分野および下位分野別に、また2つの時期(1985年～89年および1990年～94年)別に分けて計算し、分析した。すなわち、論文数、平均研究レベル(基礎から応用まで)、相対的影響度、相対的引用指標、被引用が高い論文数。

### 結果

- 研究所は、検証した2つの5ヵ年期間のあいだに論文発表を着実に 67%引き上げた。ほとんどすべての部門が論文発表を拡大したが、このうちあるとくに印象的な部門では 246%も拡大した。
- 10 部門のうち5部門が論文の 95%を占めている。
- 当該研究所は 100 万ドル当たりの特許では世界の7大政府研究所の中で1位を占めているが、100 万ドル当たりの論文では6位にすぎない。100 万ドル当たりの特許と論文では当該研究所は AT&T に匹敵する。当該研究所は検証した7つの政府研究所の中で研究開発予算が最低である。
- 研究所論文は、グループとして引用やその他の尺度についてやや平凡に見える。しかしながら、研究所は機械工学、固体物理学、工学および材料科学などの分野で世界級の研究を行って

いる。これらの分野は研究所の中核分野に関連している。全体的に低い数字は当該研究所の中核専門分野の周辺にある論文だからである。研究所が周辺領域で論文発表すれば、これらの論文の引用度が高くないのは普通である。

- 重点分野のシフトは明らかである。当該研究所はこの 10 年間物理、工学が強化され、化学が弱体化された。物理学においては、固体物理学において研究の量と影響度は拡大し、化学物理学の影響度は縮小した。

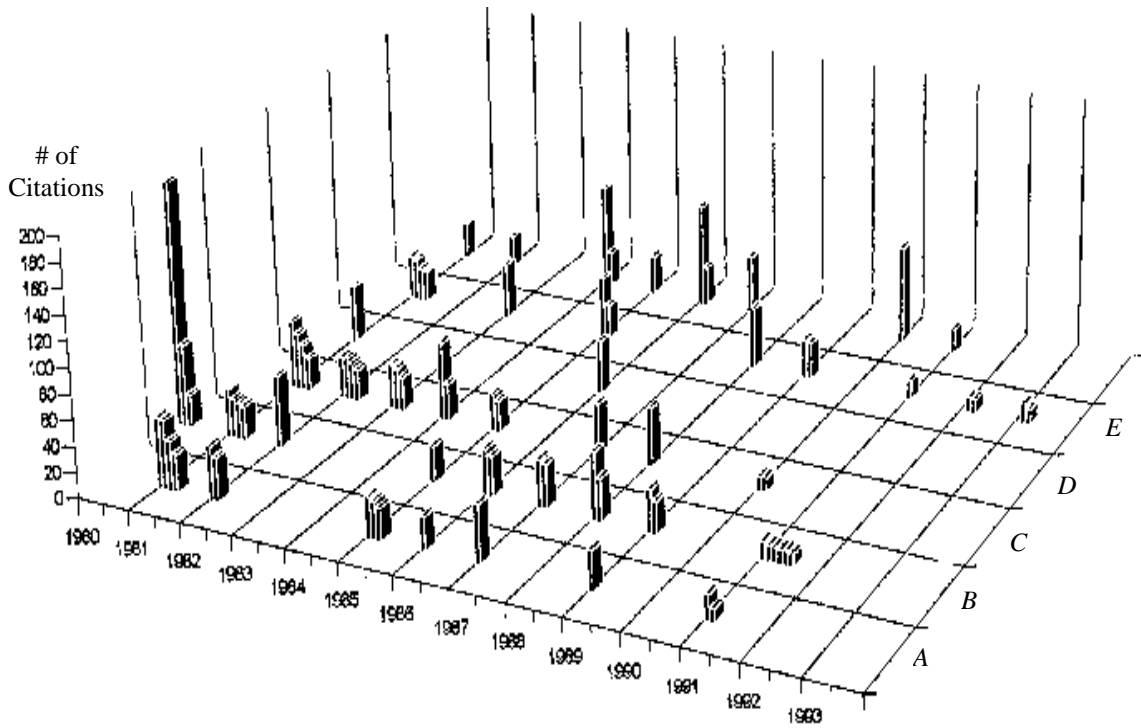
- 論文数が多い部門の各々はすぐれた分野と劣った分野を併せ持つが、これらが特定された。図4-5 は当該部門とその引用度が高い文献を示している。

#### コメント

論文発表の出版は当該研究所の主たる業務というわけではなく、この分析が研究所活動の完全な評価として意図されていたわけでもない。実際のところ、CHI はまた当該研究所の特許ポートフォリオの分析を行った。当該研究は再編成に先だって部内評価として行われたが、これにより研究所の発表論文、傾向の特定、すぐれた分野と劣った分野、およびその他の研究所との比較などに関する定量的情報が得られている。こうした情報は、包括的な評価にはならないとしても、他からは得られない洞察を与えてくれる。

图 4-5

Highly Cited Papers



\*Papers with 25+ citations 1980-90 or 10+ citations 1991-94

\*\*There are no papers with 10+ citations after 1993.

Note: Vehicle Prop., Vehicle Struct., Adv. Comp. and Survivability Directorates have no highly cited papers.

## 10. 伝統的にマイノリティの多い研究機関の研究能力向上のためのプログラム、RCMI の影響の 評価

報告者:Patrick Thomas

### 目的

RCMI(Research Center in Minority Institutions)プログラムが設定されたのは マイノリティグループ出身の学生の比率が高い学術研究機関の生物医学や行動研究の能力を向上させるためであった。RCMI の資金は 最初 1985 年に研究機関に助成された。計量文献学の評価の目的は RCMI 資金供与がこれらのマイノリティ機関が生み出した研究論文の質と量の増加につながったかどうかを評価することであった。

### 方法論

評価は少なくとも 10 年間 RCMI 資金を受けていた機関の発表論文の分析にもとづいていた。分析には2つの期間が選ばれた。第1は 1981 年～84 年で RCMI が資金供与を受ける直前の期間、そして第2は 1993 年～97 年で機関が多年にわたって資金を受けた後の時期である。機関の論文の多くの特徴が分析された。これらには論文の多くの特徴、論文を発表するジャーナルの質、論文の引用の影響および他の機関との共著論文の比率など、が含まれる。

### 結果

機関は RCMI 資金を受ける前のそれぞれの研究経験の程度によって3つのグループに分けられた。結果は以下のとおり。

- 経験が最も乏しかった機関は論文の量や影響度を拡大しなかった。しかしながら、それらの機関は共著のレベルを向上させていた。これは RCMI 資金のおかげで彼らが一般の科学コミュニティとの関係を拡大できたことを示している。

- RCMC 資金供与はそれまでの研究経験レベルが中程度の機関に最大の影響を与えた。これらの機関はより多くの論文を執筆し、それらをより質の高いジャーナルに発表し、そしてその後の研究からより多くの引用を受けていた。
- 最高レベルの研究経験をもつ機関の論文発表記録は RCMC の資金供与を受けた後もきわだった変化はしなかった。その理由は、RCMC の資金供与がそれぞれの総研究予算に占める比率が比較的小さかったからであろう。

#### コメント

従来の研究経験が研究資金供与に影響を与える可能性を立証する場合、この計量文献学の研究によって将来の資金供与プログラムを設計する上でマネージャーに有益な情報が得られた。将来、研究の成果に最大限の影響を達成するためには、この種の資金を研究経験が中程度の機関に充てるのがベストかもしれない。経験が最も少ない機関には別の種類のプログラムが必要かもしれない。

計量文献学による評価によって比較的大型の研究の一部を RCMC プログラムの影響下に置くことができた。より大型の研究には RCMC から資金供与された機関が助成金、研究スタッフや学生の数および研究支援のインフラストラクチャの開発などの獲得競争に成功した原因の分析が含まれていた。

訳注) RCMC は NIH の研究助成プログラムの一つである。

## 11. 科学および社会科学研究を結びつける学際研究機関のアウトプットの評価

報告者: Diana Hicks, January 2000

### 目的

学際的な社会科学指向の、米国政府の支援を受ける外国研究所の国際的な科学的影響を評価すること。

### 方法論

本研究における課題は、客員研究者が大きく貢献する機関の書誌データを作成することであった。科学者は当該機関を離れた後も自己の論文を執筆し続けるかもしれないので、機関の住所は論文に記載されていないかもしれない。したがって、次の3つの論文発掘方法を組合せて使った。米国の客員研究者の調査、当該研究所自身の論文データベース、論文に記載された当該研究所の住所の検索、である。上記3つの方法は突合し、データクリーニングをして、引用分析を行えるような分析的データベースを作成しなければならなかった。書籍や報告書は当該研究所の成果の重要な要素になっている。ベンチマークが利用できないため当該情報は依然として分析的というよりは記述的なままだが、しかし、これらの研究への引用が発見され、報告された。

書誌データがひとたび作成されると、標準的分野の分類や引用標準化の手続きがとられた。

### 結果

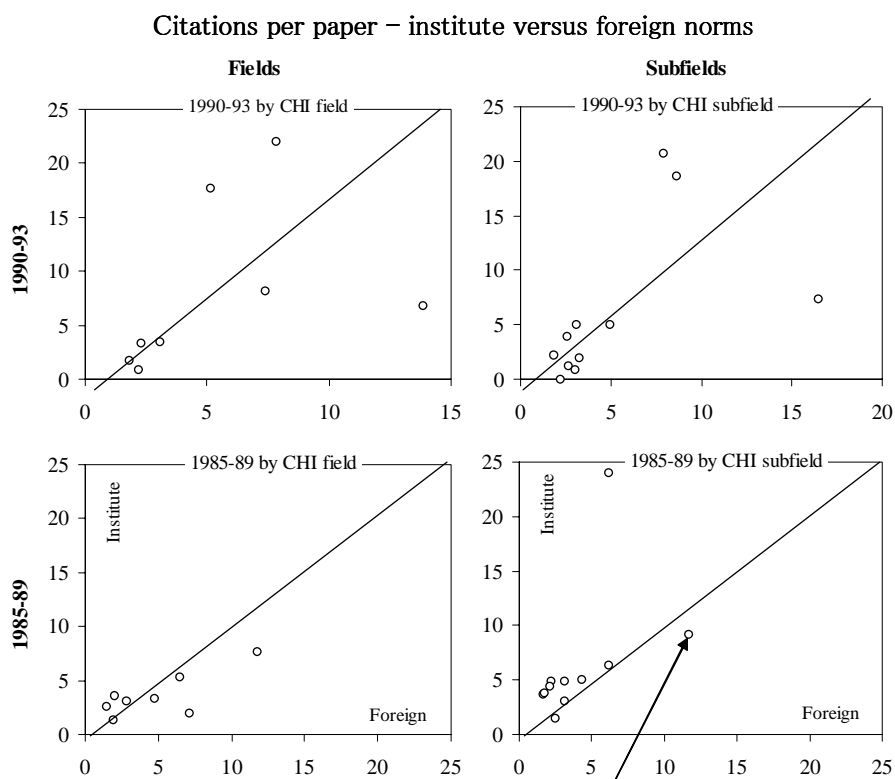
書誌データにより当該研究所の非常に独特な研究プロファイルが確認された。科学分野の論文は当該研究所の論文の 59%、社会科学分野が 41%となっており、多分組合せは非常に均衡していただろう。当該研究プロファイルはまた科学分野の中でも異例であって、地球科学(17%)、数学(8%)、およびエンジニアリング(11%)の組合せを示している。これらはより小さく、人気が高い研究分野である。このことから分かったことは、当該研究所は真の学際的研究を追及し、よく口に出されるが、ほとんど実行されない理想を実現していることである。

分野平均値と比較して引用パフォーマンスを評価したところ、いくつかの顕著な分野や非常に引用度が高い論文については国際レベルで影響を及ぼしていることがわかった。図 4-6 参照のこと。

コメント

非主流の研究所の評判は不当に低くなることがある。この外国研究所はその論文発表においてマイナーな科学分野を社会科学分野と異例の形で結合している。この結果は研究所論文の分布を他の非大学研究所の論文分布と比較すれば、もっと説得力があっただろう。定量的指標は分野と外国/国内の相違について標準化できるので、当該研究所の各分野への貢献を評価する適正な分析を行うことができよう。

図 4-6



In each of the 4 graphs, the low field represents the Institute's papers in *Science* and *Nature*. The institute contribute mathematical modelling and environmental science to these journals whose citation norms are dominated by the much more highly cited fields of biomedicine. The diagonal is roughly the line on which the Institute's and foreign norms are equal.



## 12. 計量文献学をウェブに拡張して成果(outcome)を評価することができるか？

報告者: Diana Hicks, January 2000

### 目的

ウェブにおける研究所の影響を評価できるかどうかを調べる。

### 方法論

研究所への言及状況を見るために Alta Vista でたった1回の検索をただけで、約 6,000 のヒットが得られた。これら項目の一部を吟味して見てわかったことは、項目の大半は非常に実質的なもので、当該研究所の研究の利用や研究所と一連の学界組織や政策団体とのあいだの協力を反映していることだった。

第2の操作を検証した。この操作ではウェブ上の参照パターンを使って権威のあるウェブサイト、すなわち、それ自体権威があるサイトから何回も参照を受けているサイトを特定した。これに使用したアルゴリズムには、1970年代半ばごろに Pinski と Narin が開発した CHI の影響度分析法の先駆けとなるものが含まれている。この方法は標準的な検索エンジンによる検索で見つけたページからはじまり、その後これらのページとのあいだでリンクするページを増やして行く。何回か試みた後以下のことがわかった。すなわち、正確に的をしばった検索語(例えば、「環境問題」よりむしろ「国をまたいだ大気汚染」)が見つかり、かつその語が、小冊子タイプのページとは反対に、ソフトウェアまたはデータなどのリソースを示してくれる研究所のウェブサイト上のページに関連付けられるときに最も望ましい結果が得られる。

### 結果

ウェブサイトの権威度を分析した結果2つのことがわかった。第1に当該研究所のリソースは政策分野で使用されていること、および第2に当該研究所は非常にユニークな研究所であることだった。研究所のリソースはウェブ上で参照されていることははっきりしており、したがってそれらは有力であることがわかる。研究所のリソースは国際政治分野で使用された。これは当該研究所のウェブ

ページと国際的政策立案団体からのウェブページとの共参照からも明らかである。当該研究所が世界でも独自の地位を占めていることは次の観察からも類推された。国際的環境政策立案分野における権威あるウェブのリストには研究組織はほとんどない。これらのリストにはその代わり政府機関や NGO が含まれる。

#### コメント

この種の研究所にとって幸いなのは、これらの方法が利用できるようになりつつあることである。というのはこれらの方法は当該研究所の政策影響力を確認するよう見えるが、これは以前は定量的には有効性を確認することはできなかったからである。しかしながら、標準化されていないため得られた結果が、パフォーマンスが強力か、普通かまたは弱いのかを示す証拠であるか否かを知るのはむずかしくなっている。本研究はまた、得られた結果が用いた検索句に非常に敏感に反応していることを示した。さらにまた、ウェブサイトをお互いに参照する多くのページで構成することによって結果に影響を与えることができた。最後に、サイバー計量文献学 (Cyberbibliometric) 分析のコミュニティで有名になってきたことがある。それはエンジンの取り扱い対象が毎日安定していないこと、および得られる検索特徴がしばしば変わることである。これでは高品質な評価には貢献できない。



CHIリサーチが米国連邦政府のために行った研究評価

**C.**

**特許間の分析: 研究成果 ( outcome ) の評価**

## 1. 技術移転プログラムの科学者の特許取得に対する影響の評価

報告者: Francis Narin & Kimberly Barry Stevens Hamilton, February 1994

### 目的

本研究は国立衛生研究所 (NIH) の科学者の特許取得行動に対する研究開発共同協定 (CRADA) の影響の計量文献学による評価であった。CRADA の目的は共同研究を通じて国立研究所から産業への技術移転を促進することであった。

### 方法論

CRADA を受けた科学者の特許取得記録を CRADA に関与する前後で分析し、CRADA が何らかの影響を与えているかどうかを調べた。さらに、CRADA の各科学者を、研究機関や地位ができるだけ類似している CRADA 以外の NIH 科学者と照合した。この対照群によって、CRADA に関与している科学者が各自の特許取得においてこれらの協定に関与していない科学者と異なっているかどうかの分析が可能になった。

本研究で用いられる科学者の特許データは、単純ではあるが、やや時間のかかる方法を用いて取得した。各科学者の名前は米国特許局記録における検索語として入力した。次いで名前に照合する特許は、それらのうちどれを当の科学者が発明したのかを知るために評価された。これは特許の譲受人とその主題を点検することによって達成された。

### 結果

CRADA の科学者は研究期間を通して CRADA 以外の科学者を含む対照群の平均5倍の特許を獲得した。彼らの特許はまた対照群の特許より引用度が高かった。サイクルタイム(イノベーションの速度)と科学への関連については、CRADA の科学者と CRADA 以外の科学者の特許は同じであった。

CRADA の科学者は各自の協定を受け入れた後特許取得率を向上させた。同じ期間に対照群もより多くの特許を生んだ。したがって、本研究は、CRADA 自体が科学者の特許取得行動に実質的な影響を及ぼさなかったと結論する。しかしながら、CRADA を受入れる科学者はその受入れの前後いずれにおいても特許対象の生物医学に対して異なった態度をとるようである。

#### コメント

本研究が提示するのは、技術移転の向上を目的としたプログラムの影響を評価するために計量文献学をどのように使用することができるかということである。本研究の興味ある点は、個別科学者の特許の特定のために本研究が使用する方法である。この方法(方法論の項で説明済み)は完全ではなく、同一研究分野、同一機関に名前が同一の2人の科学者が存在するというような問題にぶつかることがあるかもしれない。これは個別科学者のレベルで特許や論文データを使用することに関連する問題を浮かび上がらせる。

## 2. 共同発明者クラスターの知的財産権管理への応用

John Perko & Francis Narin, December 1994

### 目的

このプロジェクトは海軍研究所(NRL)のために実施した。その目的は NRL の知的財産権、とくにその特許管理を改善することであった。これには特許から得られる特許ライセンスの最大化、研究所内における個々の科学者の貢献の認定および NRL の発明権の保護などの問題が伴う。

### 方法論

NRL の特許は、CHI のブレイン・マッピング・ソフトウェアをつかって同機関の発明者にもとづいてグループに分類された。ブレイン・マッピングの背景にある考え方は特許の共同発明を通して結びついている発明者クラスターの特定である。ブレイン・マップによって特定されたグループ(図 4-7)はある技術における特定の問題を研究する傾向がある。ブレイン・マップはまた特定分野における研究作業の中核をなしていると見られ、かつ非常にたくさんの特許に名前を連ねている発明者を特定する。

NRL の特許ポートフォリオを共同発明者クラスターに分けた後、これらの発明者クラスターが分析され、最も引用度の高い特許が特定された。引用度の高い特許は重要な技術情報を含んでいる傾向があるとする仮定にもとづいて、これらの特許はライセンスの有力候補になると見なされた。

### 結果

ブレイン・マップは引用度が高いたくさん発明者の研究を浮かび上がらせた。これらの発明者は他の多くの組織が使用している技術の研究をしている。このことには2つの意味がある。第1はこれらの特許がライセンスの有力候補であること、および第2はこれらの発明者が自分たちの将来の研究作業を支援してもらうべきことである。

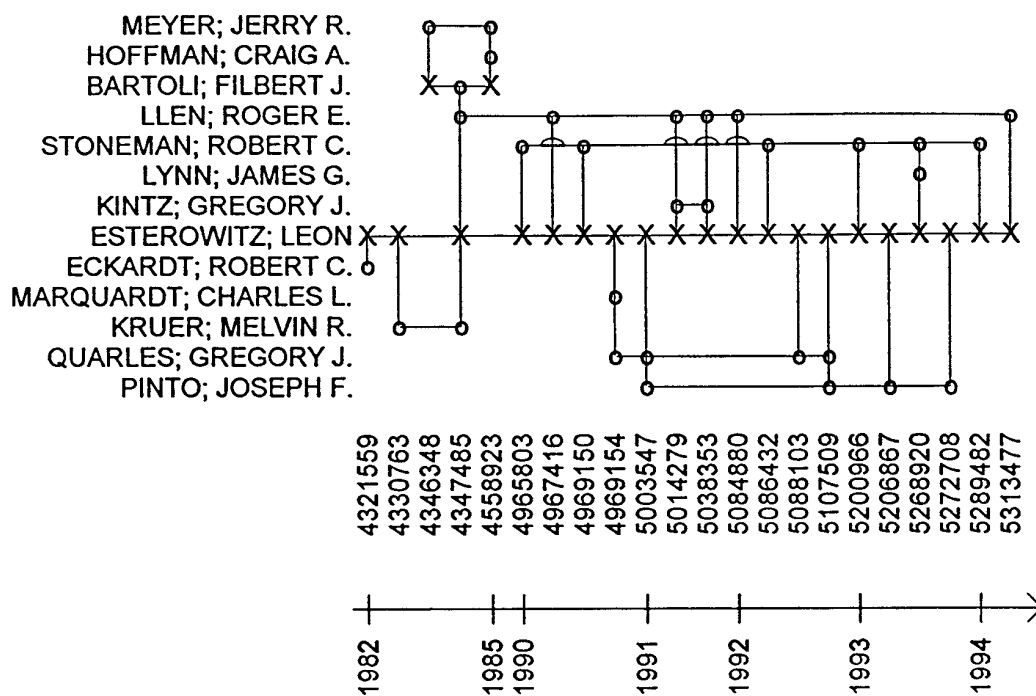
この分析でさらに明らかになったのは、NRL の一般的特許はしばしば平均以上に引用度が高いことであった。これは NRL はその後の開発に大きな影響力を持っている技術を創出していることを意味していた。NRL 特許の多くはまた、たくさんの科学文献を引用したが、それは彼らが最新の科学的発展を利用していることを示していた。

#### コメント

ブレイン・マッピングは有益な手法で、これにより組織内の主要発明者の特定、および特許ポートフォリオをさまざまな技術グループに分けることができる。それはまた特定の発明者が依然として特定の組織に残っているかどうか、または他の組織に移動したかどうかを決定するために使用できる。これは重要な情報である可能性がある。組織の大半では、特許権の取得に当たって少数の発明者に依存している場合が多いからである。

図 4-7

#### Co-Inventor Cluster





### 3. 研究所の特許ポートフォリオの評価

報告者: Anthony Breitzman, Margaret Cheney & John Perko, Decemser 1994

#### 目的

プロジェクトの目的は、陸軍研究所(ARL)の特許ポートフォリオをさまざまな主要な世界的研究所と比較することである。とくに、その目的は ARL ポートフォリオの長所と短所を特定し、それらを同じ分野で活動する研究所のポートフォリオと比較することであった。

#### 方法論

ARL の特許は、ARL 自体の手で特定された。これはすべての米軍特許を付与されるのは陸軍長官に対してであって、軍の個べつ部署ではないからである。比較グループとしては政府、大学、商業ベースおよび外国の研究所などの 14 のグループが選定された。これらには NRL、AT&T ベル研究所、スタンフォード大学およびトヨタ研究センターなどが含まれた。これらの 14 研究所の特許を特定するために譲受者の名前、発明者の名前および住所の組合せが使用された。

分析のため多くの SIC 製品グループ内の特許が選定された。これらのグループには電子部品・付属品、専門・科学器具、および非電気機械などが含まれた。これらの特許は、CHI の特許指標を使って分析され、それらのインパクト、イノベーションの速さおよび科学との関連付けなどが特定された(図 4-8)。

#### 結果

政府研究所のうち、ARL と NRL は最大数の特許を保有していた。しかしながら、CHI の特許指標によれば、NRL の特許は政府研究所の中では最も優れており、他方、ARL の特許は NRL に比べれば内容が並みである。ARL は科学的研究への結びつきが弱く、その特許は最新の科学の発展に立脚していないことを示している。

政府研究所はすべて商業ベースの研究所より特許指標が弱い傾向があった。これは政府研究所が創出した特許のより洗練された性格によるものかもしれない。

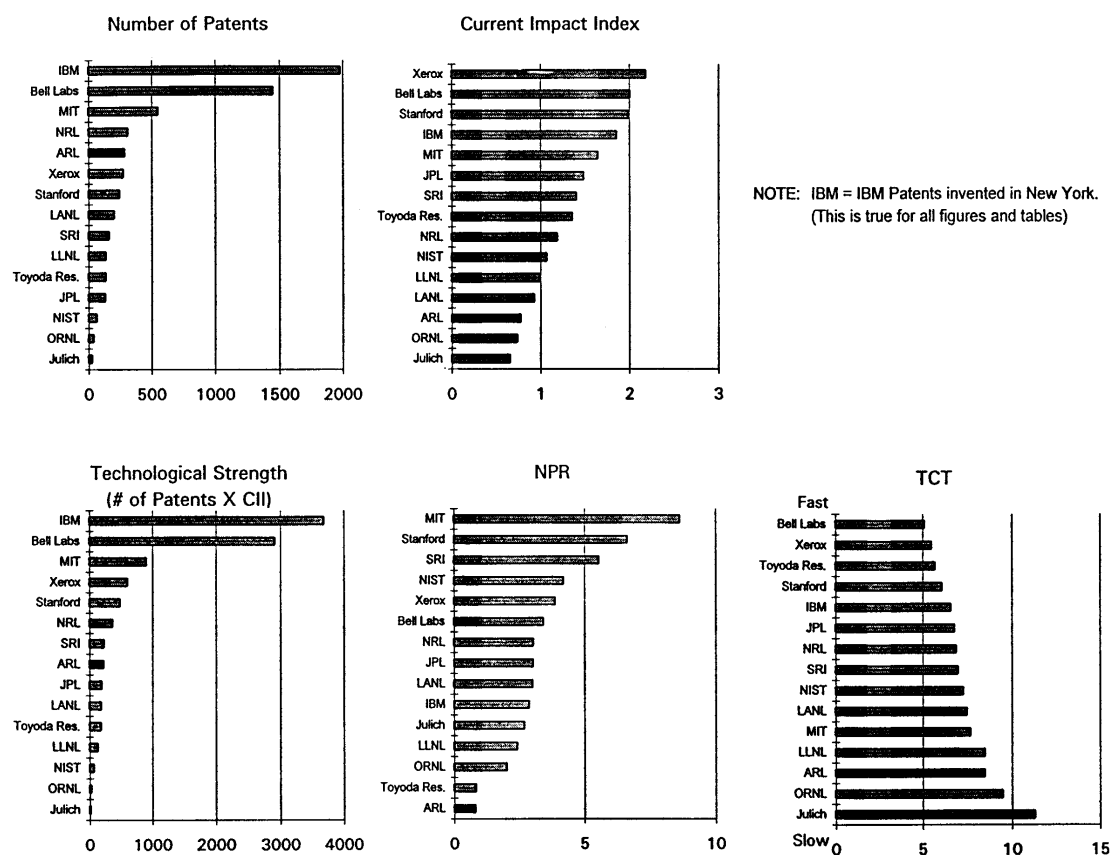
ARL についてわかった好ましい 1 つの事実は、同機関が 1 現役発明者当たり他のいずれの比較対象研究所より多くの特許を創出していることである。これは、ARL の発明者が公立、私立いずれの研究所におけるカウンターパートより生産的であることを意味する。

コメント

本研究は特許指標が特定の組織の完全な特許ポートフォリオを比較するのにどのように利用できるかを示す。それらの指標のおかげで個別特許の広範な分析を必要とせずにさまざまな組織の相対的な長所に対する定量的洞察が得られる。この種の研究の最も時間を取られる側面は、各組織についての正確な特許リストを作成する作業である。

図 4-8

Patent Activity 1990-94



#### 4. 特許ポートフォリオの引用分析による技術移転の評価と増加

John Perko, June 1995

##### 目的

本プロジェクトは米国海軍の海軍指令センター・監察センター (NRAD) の特許ポートフォリオの評価であった。その目的は NRAD の特許技術の量と質の概観を提供し、また技術移転の機会をもたらす可能性のある特許を特定することであった。

##### 方法論

NRAD は自身の特許リストを分析してもらうため CHI にそのリストを提供した。CHI は次いで自身の製品グループにもとづいて特許を標準産業分類 (SIC) に関連した 1 組のコードに分類した。製品グループの例は次の通りである。電子部品、電子式送信・配信、専門・科学器具、およびオフィス/計算/会計機。米国特許局は製品グループコードを個々の特許に割当てている。

次いで特許間の引用パターンを分析した。この分析には各カテゴリーの NRAD の特許の質と量の比較、および科学研究に密接に結びついた特許の特定が含まれた。分析にはまた他の組織からの特許によってしばしば引用されていた NRAD の技術分野の議論が含まれた。これらの技術は技術移転の候補になる可能性がある。

##### 結果

NRAD は 1980 年代を通して低いと比較的安定した率で特許を取得し、その後 1990 年以降特許取得率を拡大してきた (図 IV-9)。NRAD の特許取得件数の著しい増加が起こったのは光学、光ファイバーおよび電気機器だった。

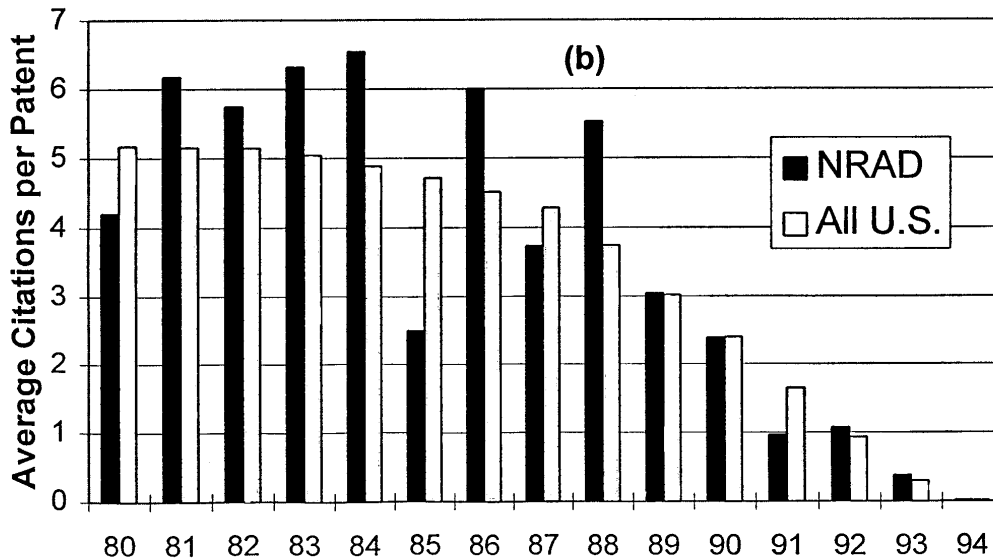
NRAD の初期の特許は同じ技術や同じ時期の他の特許に比べると引用度が高かった。しかしながら、NRAD のその後の特許は引用度が下がった。これは、NRAD が 1990 年以降もっと多くの特許を創出しはじめる一方で、これらの特許はその他の組織からそれほど興味を引かなかったことを示している。

NRAD の特許への引用件数の半数以上は、米国の民間企業から、また 3 分の 1 は外国のソース、主として民間会社からであった。これは、NRAD が民間会社との技術移転、ライセンスおよび提携の機会に恵まれたことを示している。NRAD の技術を定期的に引用する会社のリストを見るとそうした動きの潜在的な目標がうかがえた。

コメント

本研究は特許ポートフォリオの運用を支援するためにどのように引用手法を利用できるかを示す。CHI は米国海軍のその他の部署の特許ポートフォリオの同様な分析を行った。

図 4-9  
Overall NRAD Patent History 1980-94



## 5. 連邦機関が維持管理しているデータベースの研究および技術に対する貢献度の評価

報告者: Michael Albert, December 1996

### 目的

本研究では連邦機関が維持管理するデータベースの特定分野の研究開発に対する貢献度を評価した。分析したデータベースは、冷却剤および冷却混合物の熱力学的特性 (REFPROP) という標準的な参照データベースで、その維持管理者は国立標準技術研究所 (NIST) であった。検証した研究分野は CFC (クロロフルオロカーボン) の代替剤であった。

### 方法論

本研究の目的は CFC 代替剤における論文・特許に関する REFPROP データベースの影響を評価することであった。このため、当該研究分野を定義する必要があった。これはキーワードから構成される技術フィルター、特許の場合は国際特許分類を使って実現した。

関連特許の特定過程には新たな段階を付け加えた。これには、当初フィルターによって特定した特許を引用した特許のすべての検証、およびそれらのうちどれを含めるべきかの決定を含む。これらの追加特許は CFC 代替技術を基盤にし、CFC に明示的に言及していないが、次代の冷凍技術を代表する可能性がある。

次いで特許と論文を 2 つのグループに分けた。それらを産出する組織が REFPROP のデータベースを利用加入しているかどうかを基準とした。

### 結果

本研究からわかったことは、REFPROP データベースの利用者は米国特許の半数強、そして CFC 代替技術に関する科学・技術文献の半数弱を占めていたということであった。REFPROP の影響は大規模な特許取得組織・論文発表組織のあいだでとくに強かった。当該分野の特許取得組織上位 10 のうち、7組織が REFPROP の利用者(上位5組織のうち4組織を含む)であった。同

様に当該分野の上位 10 論文発表組織のうち 5 組織が利用者であった(上位 10 のうち全商業組織 3 を含む)。

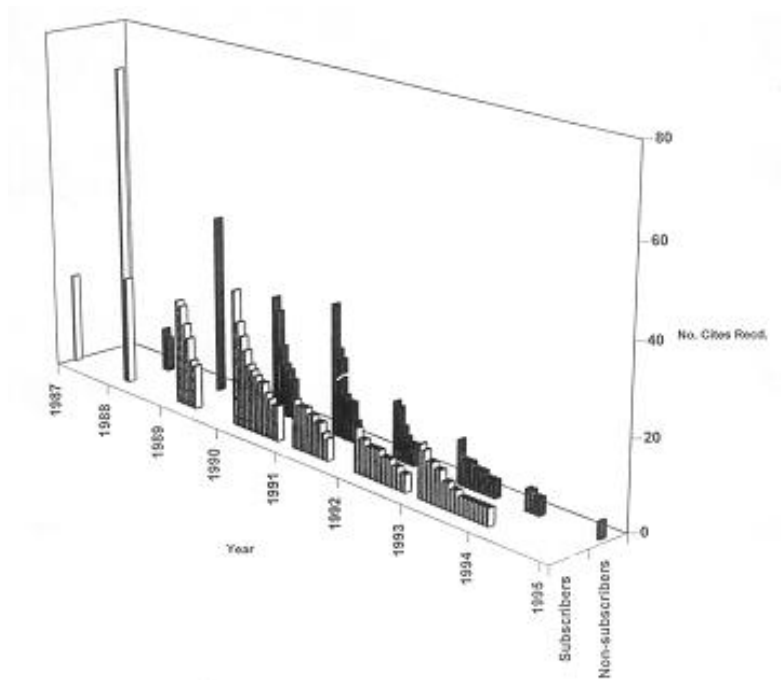
REFPROP の利用者はまた、CFC 代替物質の引用度の高い米国特許のほぼ3分の2を占めている(図 4-10)。REFPROP 利用者が当該分野における米国特許の半分以上を占めていることから、これは予想外に高い比率である。つまり、REFPROP 利用者は、平均より影響の大きい研究を生み出す傾向がある。

#### コメント

本研究は計量文献学を拡張した方法で、連邦機関が維持管理しているデータベースが研究や技術開発に及ぼす影響を評価した。データベースのようなりソースの影響は普通、計量文献学にはなじまないと見られている。しかしながら、この計量文献学の研究から連邦機関の維持管理になるデータベースがいかにして研究開発の重要な支援源になり得るかがわかった。

図 4-10

“Stick Diagram” Comparing Incidence of Subscriber and Non-Subscriber Top-Decile CFC Replacement Patents



## 6. 政府特許ポートフォリオのマネジメントのための特許引用マッピング

報告者: Anthony Breitzman, Margaret Cheney & Francis Narin, July 1998

### 目的

本研究の目的は NIH の特許ポートフォリオの運用の改善に役立つような情報の提供だった。とくに、本研究は NIH が保有する引用度の高い特許や特許群の特定に焦点を充てた。これらの特許はライセンスの候補になるかもしれない。

### 方法論

NIH は CHI に自身の特許ポートフォリオにある特許リストを提供した。CHI は次いで NIH 科学者の手を借りて、これらの特許を7つの区分に分けることにし、このため米国特許分類(POC)、標準産業分類(SIC)製品グループおよびキーワードの組合せを利用した。7つの区分とは、遺伝子工学、生物に影響を及ぼす医薬品、その他の医薬品、診断学、計測、有機化学およびその他などである。

特許分類は NIH の特許ポートフォリオの長所と短所を特定するために比較された。より詳細なレベルでは、引用度の高い個別の特許がそれらを引用する組織とともに明らかにされた。当該特許はライセンスの潜在的な候補であり、引用組織は潜在的なライセンス相手である。非常に多くの特許や引用度の高い特許を有する発明者も特定された。

### 結果

本研究の主たる研究結果は、NIH の特許が引用の影響や科学との関連などの点で生物工学や薬剤のトップ企業の特許に匹敵することである。また NIH 特許の多くは、企業にしばしば引用され、それらの特許はライセンスの機会をもたらす可能性があることを示している。

NIH の計測関連特許はとくに引用度が高かった。それらの特許は平均して他の大半のカテゴリの特許のほぼ 2 倍の引用を受けていた。

発明者の分析から明らかになったのは、NIH 特許の比較的多くが少数の科学者によるものだけということである(図 4-11 参照のこと)。多くの研究組織で生産性に関してこれと同様な傾向が報告された。これは少数の主要発明家が重要な役割を担っていることを示している。

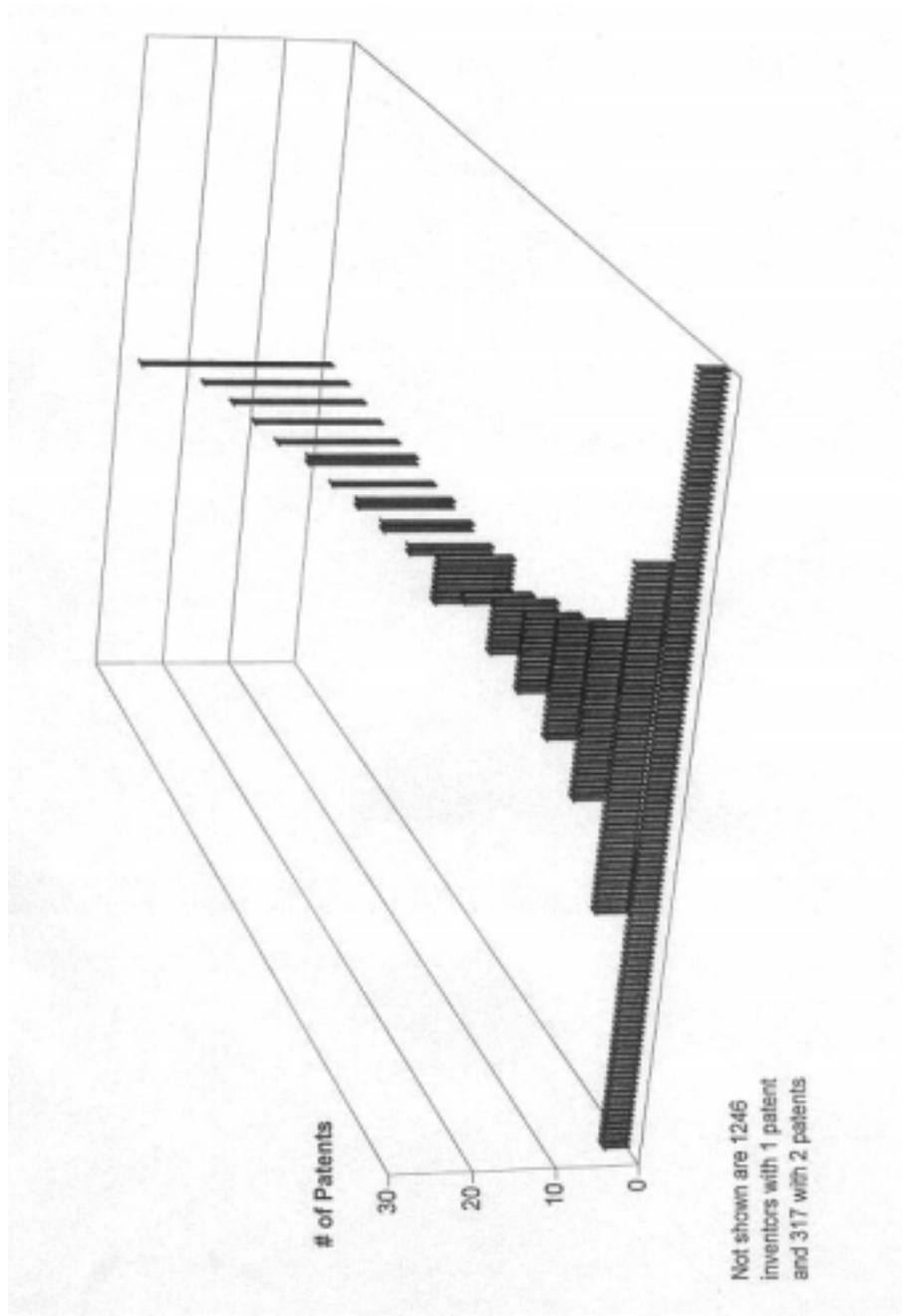
#### コメント

NIH の特許ポートフォリオなどの非常に重要なポートフォリオのマネジメントはむずかしい場合がある。本研究の結果、技術的な長所やライセンスの機会の特定のために特許ポートフォリオを分析する方法にはさまざまなものがあることが明らかになった。これらのテクニックは、追加的な収益の創出のため NIH などの連邦政府から資金供与されている団体に利用できるかもしれない。



图 4-11

Number of NIH Inventors VS. Number of Patents



CHIリサーチが米国連邦政府のために行った研究評価

D.

**特許・論文間の分析:  
研究と技術の連関 (linkage) による  
成果 (outcome) の評価**

1. 科学ベースの技術において、一国の役割は応用が進むにつれて消滅する。

**報告者: Francis Narin & Dominic Olivastro, August 1994**

関連文献: J. Anderson, N. Williams, D. Seemungal, F. Narin and D. Olivastro.

"Human Genetic Technology: Exploring the Links Between Science and Innovation,"

*Technology Analysis and Strategic Management*, 8, 2, 135-156, 1996.

### 目的

本研究では遺伝子工学やヒトゲノム技術(HGT)に関連する特許技術と関連研究の特性や英国の発明家や科学者の役割を評価した。

### 方法論

ヒトゲノム技術に関する特許セットを作成するために、USPTO 特許システムに対してフィルターが適用された。フィルターとしては IPC (国際特許分類) の 18 部類と USPOC (米国特許局分類) の 25 以上の特許を抽出し、遺伝子、遺伝学およびその他 6 つのキーワードを組合せた 8 つのキーワードを適用した。その結果抽出された 1,500 の特許を CHI と顧客が精査し、当該特許の特性や適用分野で明らかにヒト以外を対象とした特許を排除し、次のような物質の化学的変更を主とした特許を排除しようとした。すなわち、治療剤として特許を交付されたが、ヒトの分子・細胞データの情報を含まないヌクレオチドおよびヌクレオシドなどの化合物である。最終セットは 1,105 件の特許で構成された。

最終特許セットを新規性および応用分野の両方から分類する試みがなされた。新規性の分類は次の通りであった。1) 核酸と DNA/RNA の短い変更された配列、2) 蛋白質とホルモン/リンフォカイン/成長因子などに関する生成過程、3) モノクローナル抗体およびそれらの生成過程、および 4) ビールス/バクテリオファージ/ワクチン。応用分野区分は次の通りであった。1) 分析物、2) 抗菌性、3) 細胞の増殖および培養、4) がん関連、5) その他、6) 生産技法および申請のないもの。

データの取得は、HGT の特許が引用した特許や論文および HGT 特許を引用した特許などを引き出すことによって完了した。特許発明者の場所、特許に関する参照の数、論文の著者の場所、

および特許に関する資金供与への謝辞(図書館で調べた)などから本研究で使用する基本データが得られた。

## 結果

英国の影響は特許取得の世代交替が進むにつれて次第に弱まってきた。言い換えれば、おそらく技術が応用されるようになるにつれて英国の影響は姿を消して行った(図 IV-12 参照のこと)。HGT 特許が引用している第1世代において英国は論文の 6.4%、特許の 3.2%に達した。英国のシェアは HGT 特許自体では 2.8%に縮小した。HGT 特許を引用する特許においては英国のシェアは 2.5%に落ちた。

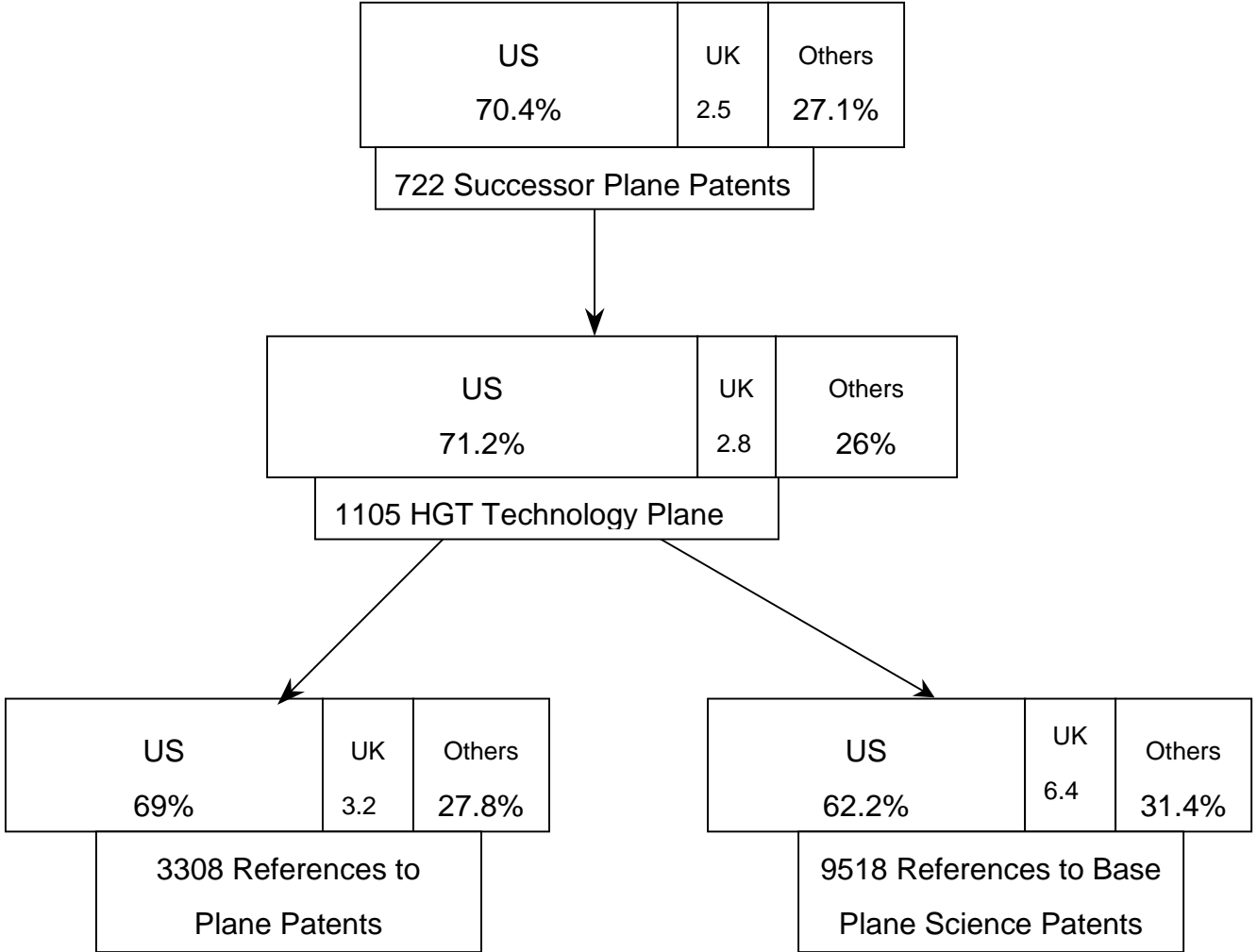
研究の結果さらにわかったことは、HGT 論文は平均して特許(3件)より論文(8件)を多く引用したこと、また発明者は自国の科学論文を過剰に引用したことである。これは科学と技術のあいだの関係では国内要素が強いことを示している。

## コメント

プロジェクトにおける最初の課題は、ヒトゲノム技術を代表する特許を決定することであった。この課題の解決にはクラスター分析・人工知能を用いることができるが、属人的な知性と知識を適用してそれを解決することも簡単である。そうするメリットは、その結果得られた特許には、必要な特許が全部、そしてそれのみ含まれるという確実性がある(高い想起力と精度)。

图 4-12

Overview of HGT 3-Plane Model



Note: Arrows indicate direction of Citation

Note: Additional 5000 Non-Patent References of various kinds

## 2. ある研究所の類似研究所に対する特許権取得のベンチマーク

Francis Narin & John Perko, August 1994

### 目的

本研究では米国における主要な 15 の研究所を検証した。その焦点はこれらの研究所が行った科学研究と研究にもとづく特許技術とのあいだの関係であった。本研究の目的は、科学研究がどのように技術開発に影響を及ぼすかを示すことであった。

### 方法論

米国特許のとびら頁にはその権利を制限する、公知項目の引用が記載されている。これらの項目は現在の特許の基礎になっている技術を示している。それらは、既存特許の参照と非特許文献の参照の2つのグループに分けられる。後者のグループの大半は科学論文で構成され、本研究の基盤になっている。

本研究は、15 研究所が発表した論文を特定するために CHI が開発した既存の科学・技術リンケージ・データベースを使用した。このデータベースを作成するために CHI は 1987 年および 88 年に発行された米国特許のすべてから非特許文献への参照 13 万件を抜き出し、そしてそれらを SCI が収録しているジャーナルと直接照合した。この方法で、当該参照のうち約 3 万は、これらの論文の米国人著者の所属機関住所も明らかにされた。

### 結果

図 4-13 は、国立衛生研究所(NIH)が発表した論文が、研究対照の 15 組織から特許による最大の引用を受けたことを示している。NIH に次ぐのは、IBM、農務省(USDA)、リンカーン研究所および海軍研究所などであった。

さまざまな研究所の優勢な分野は予想通りであった。NIH は生物医学と臨床医学で優勢であり、USDA は生物学で引用度が最も高かった。そして IBM は物理学および工学でたくさんの引用を受けた。

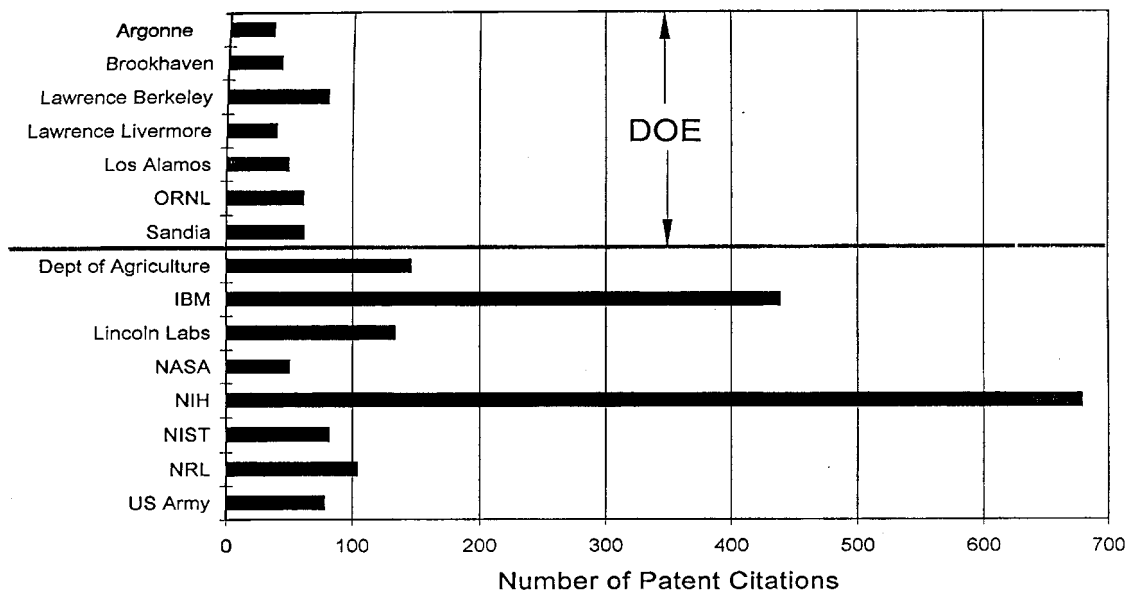
引用パタンには強い国家的な要素があった。ほとんどすべての研究所について、それらが受ける引用の 70%～90%は米国に本拠を置く組織からのものだった。すべての米国特許に占める外国のシェアは約 50%なので、米国の科学は海外の技術以上に米国の技術を促進していることがわかる。これは、発明家が好んで自国の科学を引用することを示す他の調査結果と一致する。

コメント

科学論文と特許とのあいだの引用の関連性にもとづいて、科学研究の技術開発に対する影響を数量化する方法が得られる。本研究は科学研究の実施の重要性を示すためこの種のデータがどのように利用できるかを示している。

図 4-13

1987-1988 Patent Citations to 1973-1986 Lab Papers



### 3. 特許に引用された研究所論文の検証 - 3種の報告書から

報告者: Francis Narin, Peter Kroll and Anthony Breitzman

全報告書の日付 1996年1月30日

#### 目的

3つの研究所は資金を供与した研究が特許技術に果たした貢献の研究を委託した。引用された科学とその結果としての引用する側の技術の品質と数量いずれもが関心事であった。

#### 方法論

著者のうち少なくとも1名が米国人であり、かつ1987年～88年または1993年～94年に交付された米国特許が引用している論文すべてのデータセットを選び、それらを大学図書館で検証し、謝辞が記されている資金援助および著者機関を探した。そして3つの機関のいずれかが資金供与または機関提携を通じて提携した論文を抽出し、さらなる分析を行った。特許や論文の指標や時間的趨勢および研究所内の著作と研究所外の著作との相違を研究した。

#### 結果

当該研究所が支援した論文を引用する特許の点数は、1987年～88年および1993年～94年のあいだにほぼ3倍に増加し、また特許に引用された個別論文の点数も同様だった(図4-14)。引用総数はさらに増加した。上位に引用された著者の研究所はその所内または所外の中心や特許に最も引用されている研究などを示していた。大学と研究所および民間法人両方で構成される上位特許権譲受人は、研究所関連の研究から最も利益をあげている会社や研究組織を示している。強力な所内プログラムを有する研究所は相当高いレベルで特許を取得していた。

特許による引用度が最も高い研究所関連論文を特定することで特定の研究主題や著者の重要性を明らかにした。論文を下位分野やレベルおよび影響特性により分類し、研究品質の指標を求めた。その他の資金に対する謝辞を一覧表にして共同資金供与源を示すようにした。



特許はまた特許権の譲受人の部門や国別に特徴を示し、そしてその後の特許による引用が最も高くなった特許を潜在的に重要な技術進歩として特定した。これらを当初の研究所支援研究まで追跡できるようにした。

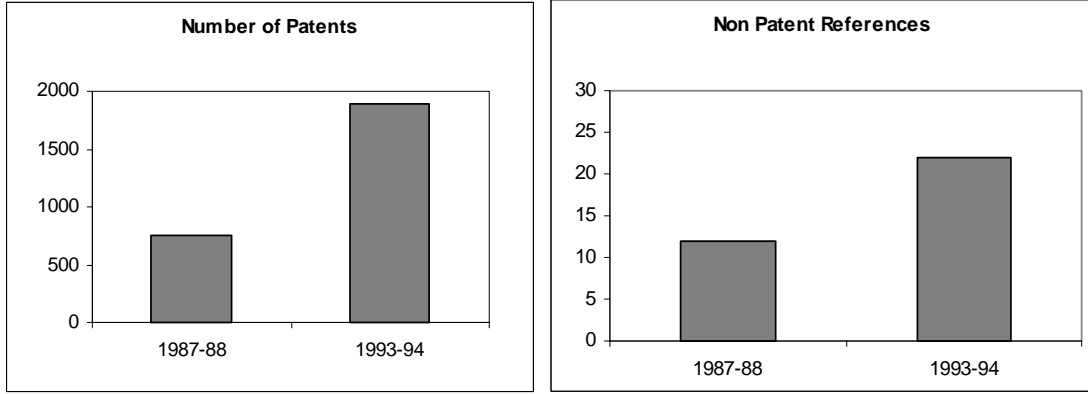
#### コメント

この研究の大きな成果は、特許・論文リンケージのデータベースを作り上げ、技術にいたる研究資金供与の多段階関係のさまざまな側面からの分析に役立つようにしたことだった。

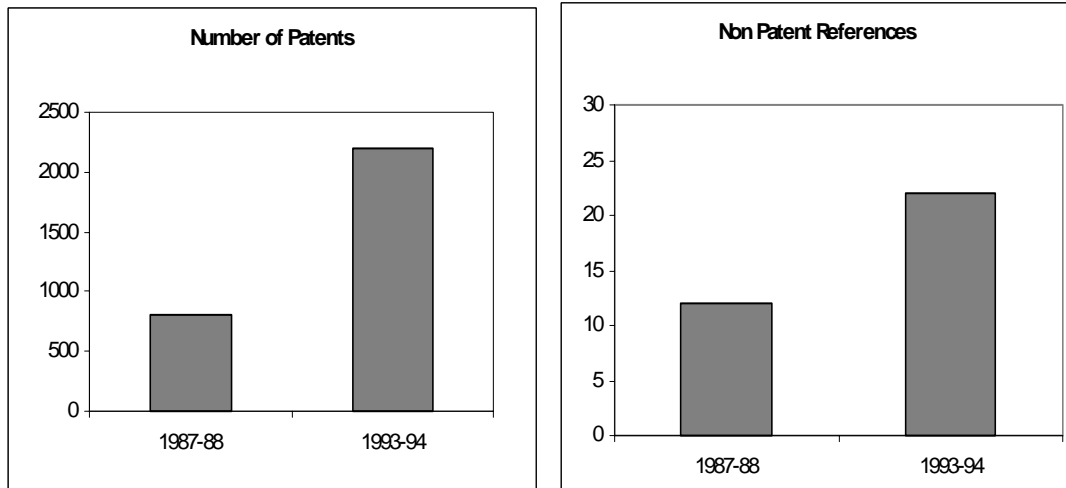
图 4-14

Indicators for 1987-88 & 1993-94 Patents Citing Funded/ Authored Papers

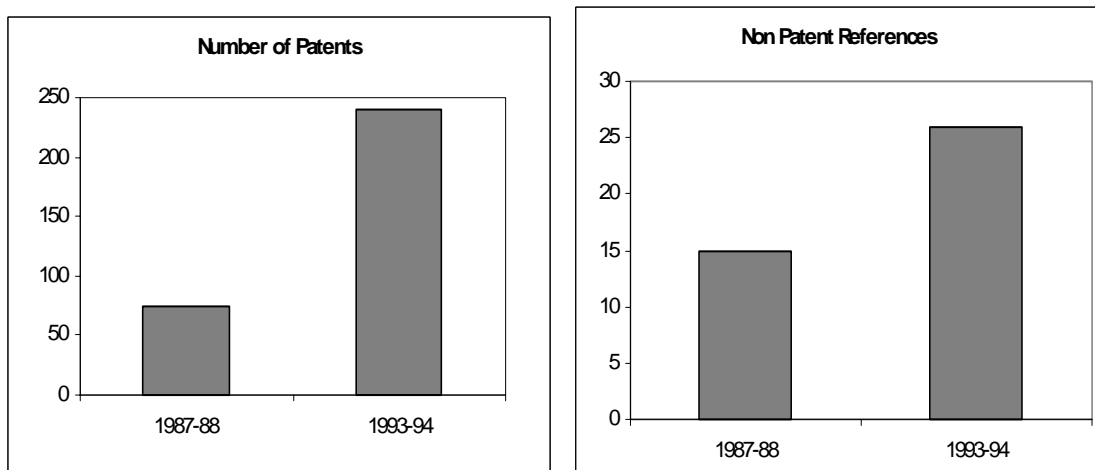
NIGMS



NCI



NIMH



#### 4. 政府機関の科学研究を参照する特許技術の評価

報告者: John Perko, October 1996

##### 目的

本研究は農業研究サービス(ARS)による特許化された工業技術への貢献を検証した。

##### 方法論

以下の2つの基準を満たす論文を特定した。1) ARS 研究所を起源とするか、米国農務省 (USDA) が資金を供与している研究、2) 1987 年～88 年ならびに 1993 年～94 年の両期間に米国特許に引用されたもの。

##### 結果

論文数および引用数は 1987 年／88 年と 1993 年／94 年のあいだに3倍に増加した。USDA 支援の論文の約半数は ARS 研究所、そして半数は大学からのものだった。ARS 論文がより多く使用され、また農業指向であり、そして大学の論文はより基礎的で、生物医学関連のものが多かった。これらの論文を引用する特許は、非常に多くの科学論文、平均 10 件以上の科学論文を引用した。これは米国における科学論文への引用が1特許当たり平均1件の割合にひけをとらない。当時の米国特許の約半数は外国所有のものであったが、ARS 文献を引用する特許の 90%以上は米国で発明されたもので、これは ARS の科学研究が米国特許に貢献している事実を示している。

1988 年に刊行され、1995 年までに発行された特許に引用された論文はすべて検証された。生物学の分野では特許に最も引用された研究所は ARS であり、3倍の被引用論文を有していた。

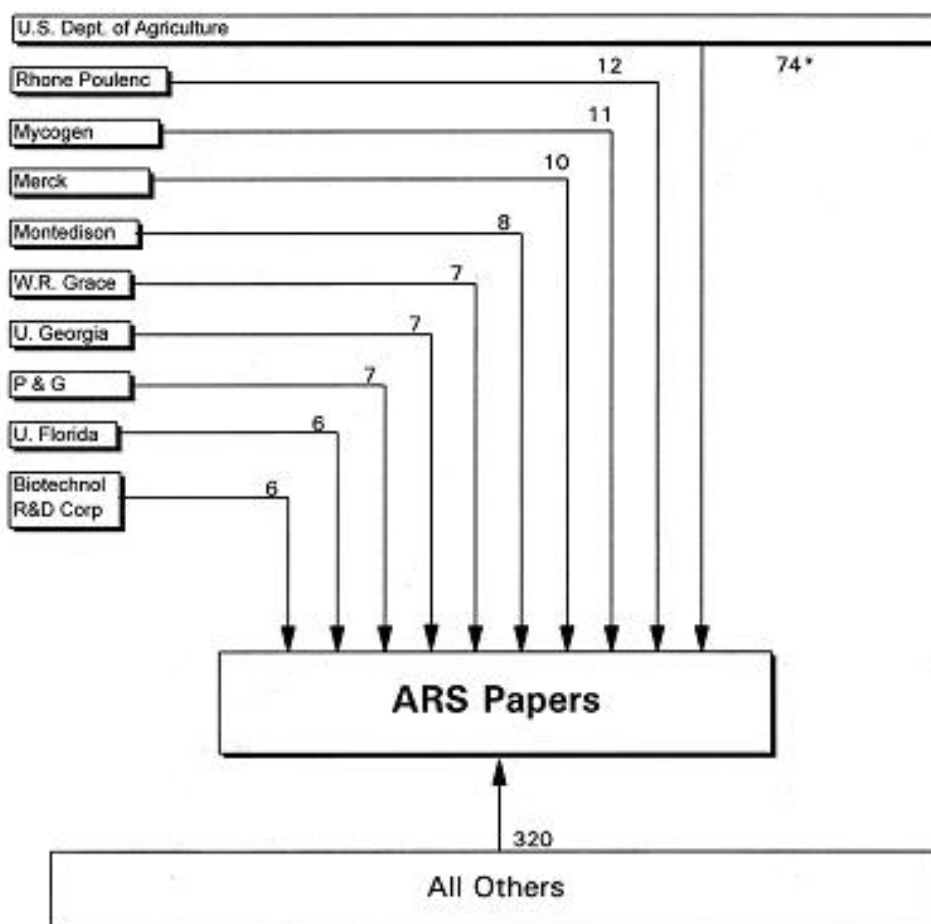
ARS 論文を参照した特許を発明したトップ組織も特定された(図 4-15 参照のこと)。最も頻繁に ARS 論文を参照している特許の製品グループも特定された。図 4-16 参照のこと。

##### コメント

CHI は類似の研究をたくさん行った。特許論文間の計量文献学は、研究の成果 (outcome) の評価に貴重な貢献をしている(これは研究結果の output とは対照的である)。論文間は計数表示なので理解しやすく、また効果の評価する上で仮定の質問に対する主観的回答の形をとっていない。この利点の一方で、特定された研究用途をドル表示できないという問題点がある。

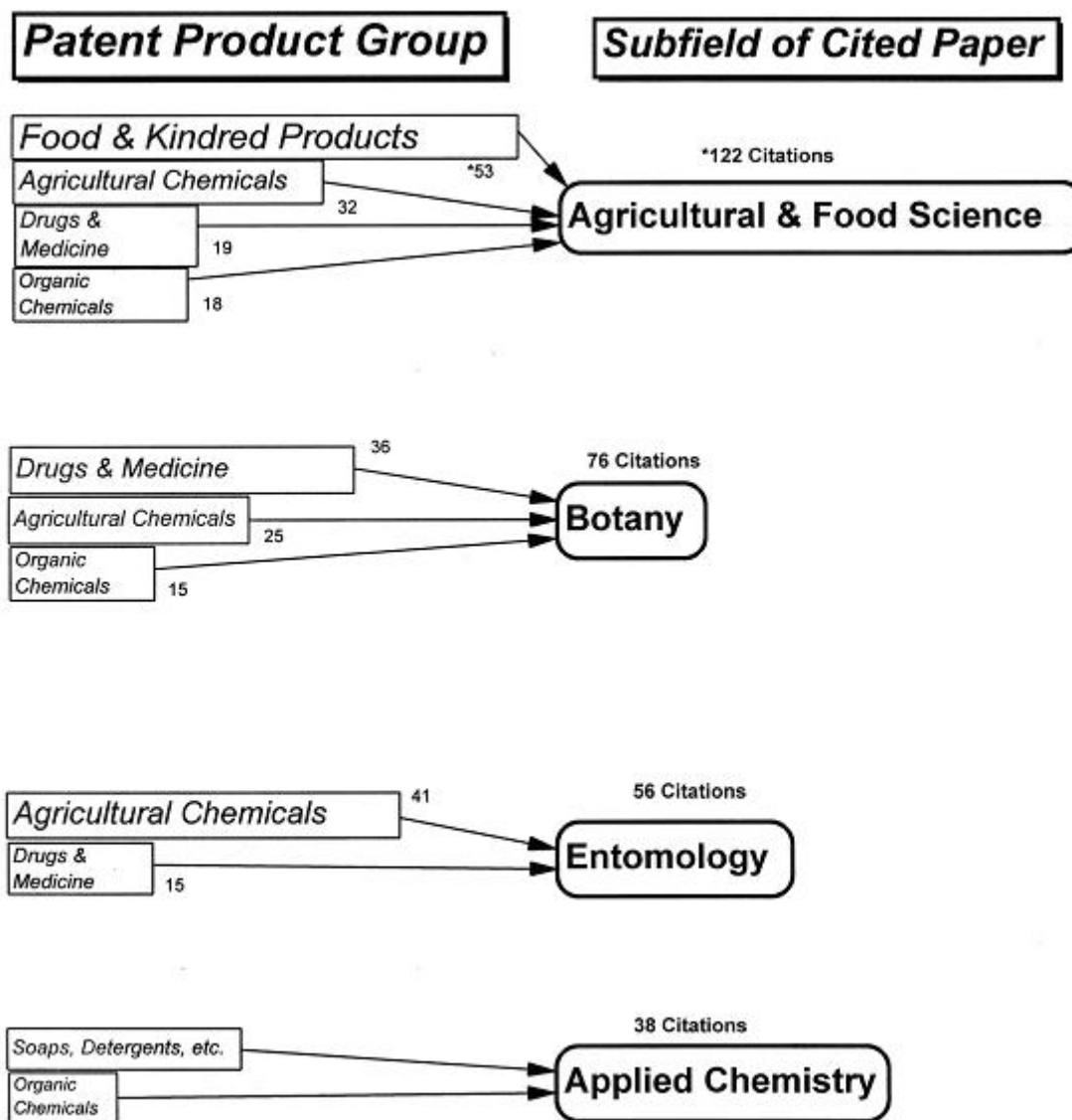
図 4-15

Top Ten Patent Assignees Citing ARS Papers



\*Counts of citations from 1987-88 and 1993-94 U.S. patents

Most Frequent Citation Connections  
Between Patent Product Groups and ARS Publishing Fields



\*Counts of citations from 1987-88 and 1993-94 U.S. patents

## 5. 眼科技術に対する NEI 資金供与研究の重要性の検証

**Leon B. Ellwein, Peter Kroll, and Francis Narin, "Linkage Between Research Sponsorship and Patented Eye-Care Technology," Investigative Ophthalmology & Visual Science, vol. 37, no. 12, November 1996, 2495-2503.**

### 目的

国立眼科研究所(The National Eye Institute)は眼科技術の状態の評価を希求し、このため 1975 年～96 年以降の特許取得、とくに NEI が資金を供与した特許技術に関する研究の判定の調査を実施した。

### 方法論

1975 年～94 年米国特許から 8,163 件の眼科技術特許を特定した。このために特許局分類にもとづいた定義フィルターおよびキーワードを使用した。それらの特許が引用している科学研究への参照を、著者所属研究所や資金供与組織の記録とともに大学図書館で調べた。本研究では時間的趨勢、研究所の著作および特許のその他の諸様相を評価したが、ここで関心のある分野は被引用論文における研究資金供与である。

### 結果

表 4-4A は研究資金源の区分(NEI、その他の NIH、およびその他、の広範な区分)別の特許数、科学参照数、被引用論分数および資金提供への謝辞などを記載したものである。参照が見つかった特許 1,192 件のうち 739 件(62%)は、少なくとも1点の論文を参照しており、そしてこの論文では何らかの形式の外部支援が記載されていた。注目する単位として特許の代わりに科学参照の記載を分析すると、見つかった科学参照の合計 4,889 件のうち 2,588 件(53%)は複数の資金供与源に対する謝辞が記載された論文を参照していた。NEI は単一研究所として 371 件(31%)の特許を有して主要な役割を果たしている。またこれら特許は NEI の支援に対する謝辞のある 606 点(18%)の論文に対する 1,001 件(20%)の参照を含んでいる。特定の支援源が記載されて

いる特許、参照および論文にかぎって見れば、NEI はそれぞれ、50%、39%、および 34%を占めている。NEI は合計 6,130 件の資金援助についての記載のうち、28%を占める。NEI がその他の NIH(すなわち、NIH は全体として見る)と合併する場合は、総パーセンテージは特許の 41%に達する。そしてこれらの特許には NIH に関連した次の参照が含まれている。被引用論文の 41%、科学参照の 34%、被引用論文の 33%および資金援助についての記載件数の 51%。米国の民間非営利部門は全体として、資金源として NEI に匹敵する。

表 4-4B は上記の科学参照の分布を著者所属研究所や資金源別に分類して示したものである。NEI と NIH は全研究ポートフォリオの 34%という大きな部分を支援しているが、その支援は米国の大学関係部門における研究に関してはとくに強力である(59%)。

#### コメント

リンケージの特定方法は比較的保守的である。限定した特許のセットから開始している。当該分野に関する研究ツール特許は、特許部類やキーワードが厳密な分類フィルターの定義に該当しない場合は、眼科技術に関するものとして選択されなかったかもしれない。

**表 4-4A. Patents with Science References, Papers, and Acknowledgments by Support Source Sector for Eye Technology Patents: 1975-1994**

<i>Support Source</i>	<i>Number of Citing</i>			
	Patents with Science References	Number of Science References	Number of Cited Papers	Number of Support Acknowledgments in Cited Papers
National Eye Institute*	371	1001	606	1687
Other NIH*	246	790	607	1452
Other US government	250	407	300	501
US private nonprofit	389	892	625	1278
US foreign for-profit	230	324	225	376
Foreign government and nonprofit	268	542	412	798
Not fully identified	32	36	29	38
Any acknowledgement	739	2588	1808	6130
No acknowledgment	453	2301	1616	NA
Total	1192	4889	3424	NA

NA = not available.

\*Combine NEI and Other NIH into a single category results in the following counts: patents = 491; references = 1674; papers = 1133; acknowledgments = 3139. For patents, references, and papers, these counts are less than the sum of the two rows in the table because of the avoidance of double counting when a single paper acknowledges both an NEI and a non-NEI NIH institute.

**表 4-4B. Science Reference Counts Disaggregated by Author Institution and Support Source for Eye Technology Patents: 1975-1994**

Funding Support Source	<i>Author Institution</i>							Total	No Information
	US Academic	Foreign Academic	US for Profit	Foreign for Profit	Federal Government	Other US	Other Foreign		
NEI*	726.2	41.7	9.2	2.0	22.5	189.7	6.8	998	3
Other NIH*	565.0	27.3	25.7	5.0	38.7	118.0	9.3	789	1
Other US government	289.1	9.3	15.2	3.0	29.1	55.2	2.1	403	4
US nonprofit	600.7	41.7	17.0	1.0	34.3	186.0	10.3	891	1
US + foreign profit	182.2	40.2	28.2	1.0	6.8	53.5	8.2	320	4
Foreign government nonprofit	55.7	355.3	22.2	13.8	5.3	16.3	71.3	540	2
Not sufficiently identified	17.5	9.5	0.0	0.0	1.0	6.5	1.5	36	0
Any support acknowledgment	1506.4	419.5	91.3	19.8	80.7	374.1	86.1	2578	10
References with no support identified	511.1	607.3	414.2	105.2	164.4	229.5	238.4	2270	31

\*Combining NEI and Other NIH into a single category results in the following counts: US academic = 1198.7; foreign academic = 67.0; US for profit = 34.3; foreign for profit = 6.5; federal government = 58.2; other US = 289.2; other foreign = 16.2; total = 1670; no information = 4. These counts are less than the sum of the two rows in the table because of the avoidance of double counting when a single paper acknowledges both an NEI and a non-NEI NIH institute.



## 6. 特許に引用された政府機関資金による論文の特徴の検証

報告者: John Perko & Francis Narin, December 1996

### 目的

政府機関資金による科学研究の特許に対する影響の検証と国内特許が政府機関支援研究を引用している程度の評価。

### 方法論

米国人著者を少なくとも1名含む、かつ 1993 年～94 年に交付された米国特許から引用されたすべての論文のデータセットから NIH が支援し、または NIH で生み出された論文を特定した。1993 年～94 年に発行された米国特許 22 万 3,000 件のうち、米国で著作された論文(11 年ウィンドウ内で引用された)1 万 3,000 点から選んだ NIH 支援のもの 500 件を図書館で調べた。特許は上記論文を約 2 万件引用し、3万 8,000 件は NIH が資金供与したものであった(NIH のものも、そうでないものもある)。特許と論文の指標が作成された。

### 結果

NIH が支援した論文や引用の約 80%は純粹に研究所外のものであった。これらの論文の 90%以上が生物医学と臨床医学のジャーナルに記載された。NIH 支援の生物医学論文をこれらの特許が引用した全臨床医学論文と比べると、当該論文のほぼ 3 分の1が NIH の何らかの支援を受けていた。少なくとも若干の NIH の支援を受けた論文は臨床医学論文に対する特許の引用総件数の約 53%を占めた(図 4-17)。引用件数および被引用論文両者を併せた件数の約 63%は大学、カレッジおよび医学部からのものであった。NIH 支援論文は非常に基本的な科学を扱っている。発表済みの生物医学論文全体と比べた場合はとくにそうである。これは応用科学ではなく、基礎科学が生物医学のイノベーションを生むことを示唆している。公共資金供与に大きく依存しているのは基礎科学である。NIH 支援の被引用論文はまた非常に有力なジャーナルに発表されている。さらに NIH 支援の生物医学論文は世界で発表された生物医学論文全体の 2 倍以上特許

から引用されている。NIH 支援の科学は特許技術において生物医学全体よりはるかに有力である

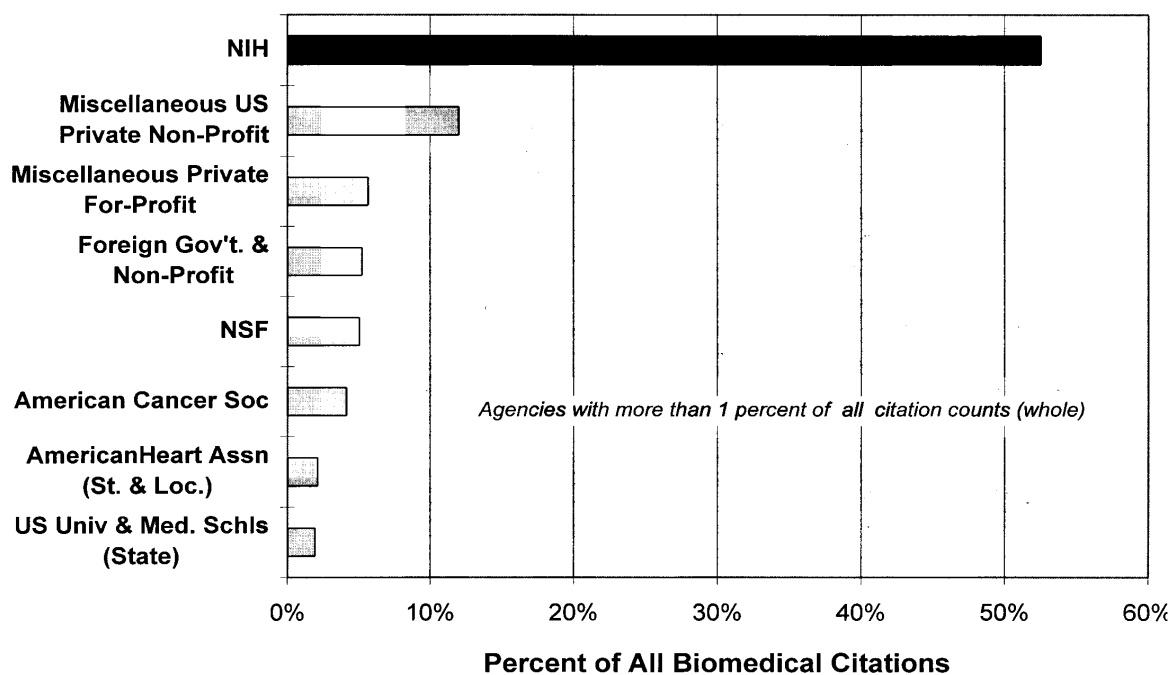
最後に当該特許の 70%以上は米国人の発明家を雇用する米国組織が保有している。NIH が支援する科学は国内関係者の手元にある技術に刺激を与えている。

#### コメント

NIH 支援の科学および非生物医学特許研究が本研究と平行して行われた。

図 4-17

Top Funding Agencies Acknowledged on All Biomedical Papers Cited by Patents  
1993-94 Patents Citing U.S.-Authored Biomedical Papers Published in 1981-91



Includes only papers with explicit support acknowledgements

## 7. 最近承認された医薬品における NIH が資金供与した研究の重要性の検証

Peter Kroll & Francis Narin, January 1997

### 目的

本事例研究は、重要技術の基盤になっている科学を特定するために、さらには研究に資金を提供している機関を特定するために、重要を再評価した一連の技術の研究の一つであった。ここでは商業ベースで重要な医療技術、例えば、米国食品医薬品局が最近承認した医薬品などに関する技術を評価した。これらの医薬品は分析の対象として選択したが、その理由はそれらがもっと古い、すでに承認済みの医薬品とちょうどその特許が発行されたばかりの将来の医薬品とのあいだの製品開発における重要な中間点を示しているからである。本研究の目的は特許情報を使用してこの医薬品のセットをそれらの基礎科学研究基盤(および研究のための資金的支援)まで立ち返って追求することであった。医薬品に関して入手できる情報を使って「トップダウン」の方法で、それらの技術を保護する特許を決定するようにし、そうした特許が引用する論文の出所をそれぞれの当初資金供与源にさかのぼって追跡した。この方法をとれば、あらかじめ各医薬品の基礎研究の発展の足跡を初めから終わりまで知らなければならないという問題や、各医薬品開発の論文調査を実施しなければならないという問題を避けることができる。

### 方法論

各分野の専門家と協力して医薬品のセットを特定した。そのセットは 1995 年から 1996 年にかけて承認された医薬品であり、"New Drug Approvals in 1995" (Pharmaceutical Research and Manufacturers of America, January 1996) および FDA のウェブページ "Approved Drug Products with Therapeutic Equivalence Evaluations" (1996 年 9 月ごろまで掲載) に編纂されたものである。

新規承認医薬品に関連した特許は次のソースから決定された。すなわち、各医薬品の FDA への承認申請書に記載された特許、メルク社医薬品索引 (Merck Index) および IMS 世界特許国際データベースである。本研究の総合調査的性格から、各医薬品の科学的、技術的論文の詳細な調査はむずかしいが、追加の特許や論文は特定できるようである。

その選択過程の結果(正確に)100 件の独自の特許が得られた。これらの特許のうち 5 件は 1975 年以前に発行されたが、対象から外された(当該日から米国特許のデータベースは機械可読形式になった)。残りの特許 95 件は当初セットの医薬品 55 点のうち 48 点を占めた。これらのうち医薬品 26 点は科学論文を引用した特許に関連していた。それらの特許には、資金供与および著作目録情報を探すために大学図書館で調べた 307 件の非特許項目が引用されていた。見つかった特許および 117 点の論文に関するリレーショナル・データベースが開発された。(155 点の科学参照が提示されている)。

## 結果

資金供与源トップ 4 のうち3つは NIH 研究所であることがわかった。科学論文を引用する特許に該当する医薬品 26 点について、図 4-18A は当該医薬品のうち 12 点は NIH の支援を記載した論文を引用する特許をもっていることを示している。117 点の論文のうち 24%には NIH 資金供与への謝辞が記載されていた。NIH の資金供与とする記述がないが、NIH 所内科学者が作成した論文の 4.5%に加えて、見つかった論文の 29%が NIH 支援であることがわかる。

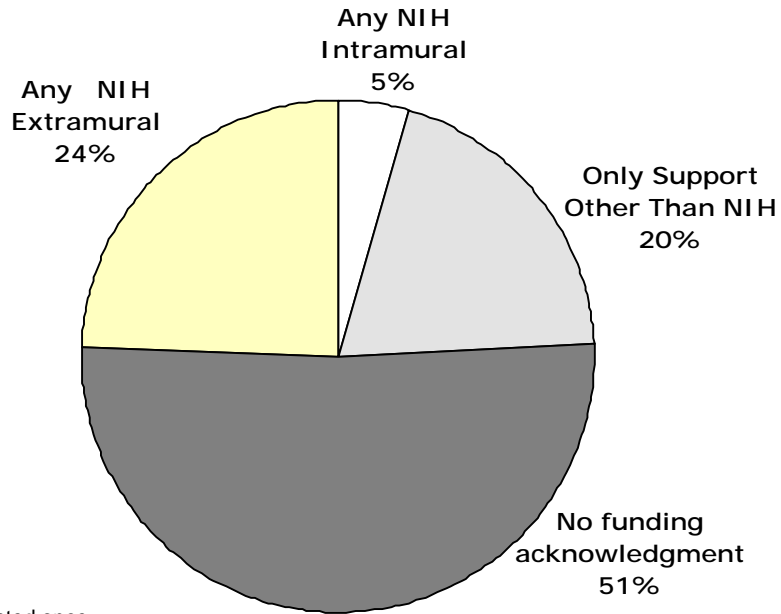
統計を調べるもう 1 つの方法は、図 4-18B に記載した引用特許をチェックすることである。ジャーナル論文の引用を含む 44 の特許のうち、18 は少なくとも1件の NIH 資金供与または NIH 論文を引用してあった。

## コメント

CHI リサーチが他の研究から見いだしたところでは、引用された非特許参照の数は急激に増加した。1975 年～82 年以降の当該セットの特許は、1特許当たり平均 0.81 の非特許参照(non-patent reference)を含んでいた。1990 年～96 年には平均値は 4.89 だった。したがって、初期の特許では関連研究の資金供与源をほとんど追跡できなかった。

☒ 4-18A

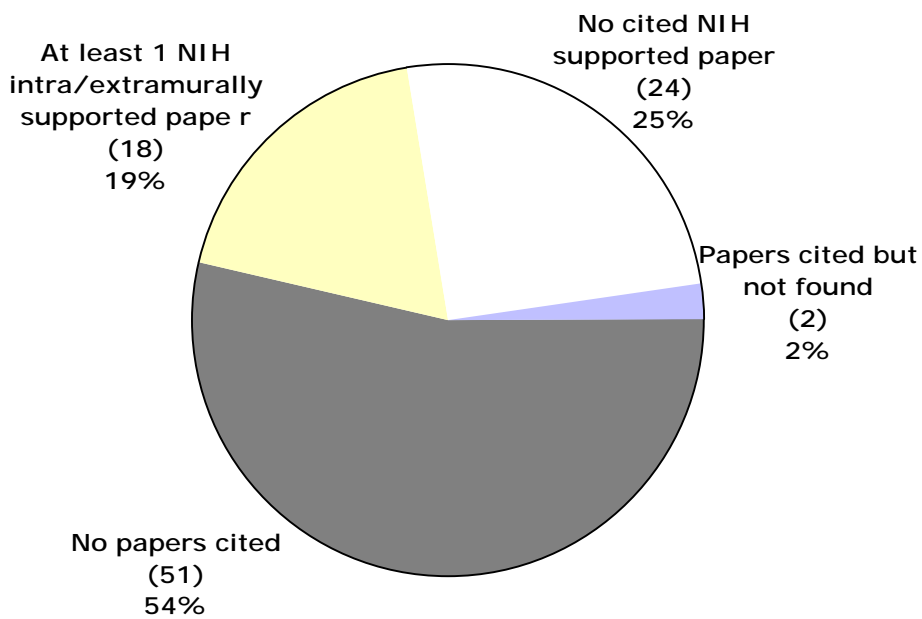
Support for papers Cited in “FDA Set” Patents\*



\*Each paper counted once.

☒ 4-18B

NIH Funding Support for Papers Cited in “FDA Set” Patents: Counted by Patent



## 8. 中核分野外における政府機関支援科学研究と特許技術とのあいだの関連性

John Perko & Francis Narin, January 1997

### 目的

NIH(国立衛生研究所) 資金供与を受けた科学研究が生物医学以外の技術に果たす貢献度の測定。

### 方法論

先行研究では 1993 年～94 年に発行された米国特許のうち、NIH の支援であることが記載されているか、少なくとも1つの NIH の住所が記載してあるジャーナル論文をとびら頁で参照している特許が特定された。本研究では特許セットからすべての医療特許を除外してある。ほとんどの医学特許を特定するのに USPTO 分類を使用し、そして残りの特許の標題と抄訳を精査して他の医学関係技術を除いた。344 の異なる論文を引用する 242 の非医学特許がこの分析の基盤を形成した。

### 結果

植物科学、病虫害防除剤、器械および計測が非医学特許の約 60%を占める。食品科学とコンピュータおよび数値手法はそれぞれ 10～12%を占めた。その他 NIH 研究を引用する分野は材料および構造、化粧品、危険廃棄物および家庭用繊維・クリーニングなどであった。引用の約 75%は、NIH 研究の中核になっている生物医学研究および臨床医学における論文に対して行われている。他の科学分野もまた引用された(図 4-19 参照のこと)。

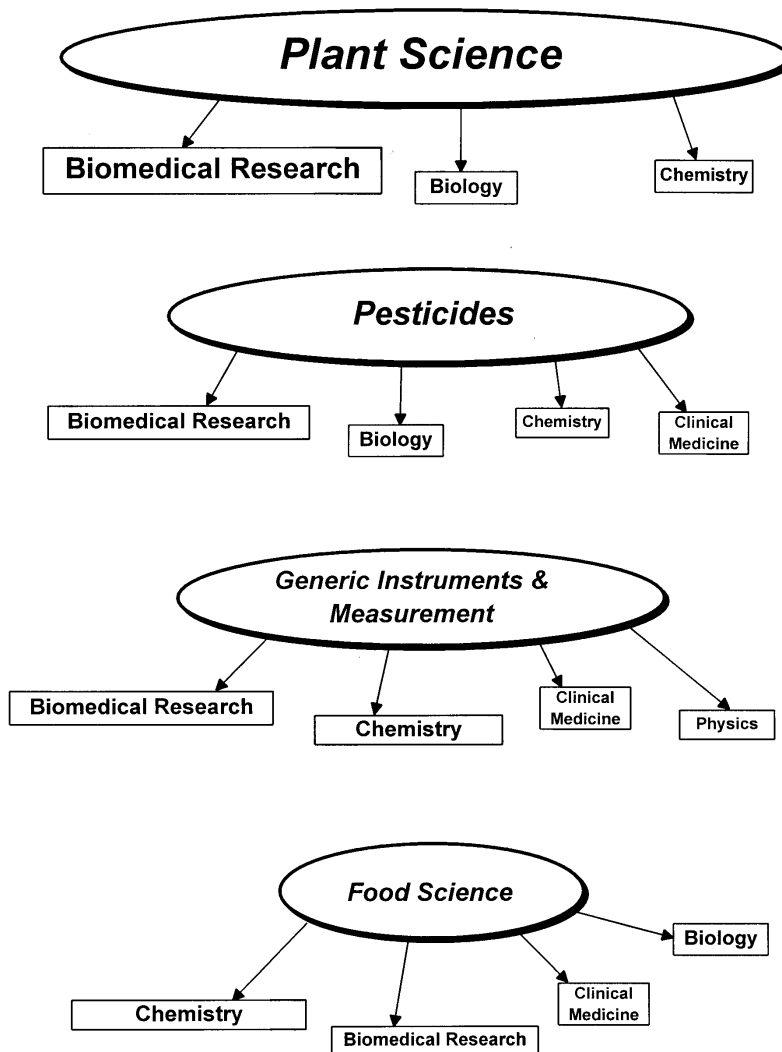
非医学特許が引用する論文は 1993 年～94 年の特許で引用されたすべての NIH 支援論文と同等か、それよりやや基礎的な科学関連のものである。引用特許の 80%は米国の組織が保有している。このことは NIH が資金供与する科学は主に国内技術で利用される傾向があることを示している。

コメント

理論上わかることは、科学研究は広範で予想もつかない影響力をもっているということである。本研究は NIH の科学がその中核にしている医学分野以外に影響力をもっている技術をマッピングしてある。マッピングの対象は大部分、NIH の中核の生物医学であって、周辺の科学ではない。政治家が科学や外国技術の便益に資金供与している点を懸念している場合は、政府機関の科学に対する国内ユーザーの大きさを立証することが大切である。

図 4-19

Non-Medical Patent Technologies Depending on NIH Science



## 9. 科学的基盤を NEI 資金供与研究にたどれる技術の経済的影響の検証

Peter Kroll & Francis Narin, December 1997

### 目的

選択した 2 つの技術に国立眼科研究所の資金供与研究が与えた影響の評価。

1 つの技術についてここに要約する。

### 方法論

医薬品ラタナプロストはプロスタグランジン基剤の、眼圧を下げるための治療薬である。初期の研究では 1975 年～94 年のあいだに 35 の特許が特定され、それら特許が引用した 77 点の論文が図書館で見つかった。これらのうち、31 点は資金供与支援が記載されていない。残りの 46 文献のうち、33 点は国立眼科研究の支援を記載しており、さらに 4 点は他の NIH ソースによる支援を挙げ、残りの 9 点は NIH 資金供与とは別の支援を受けていた。

### 結果

図 4-20A は当該分野の特許と論文とのあいだの関連を示している。特許はページの上部に、また論文は下部に示してある。NEI の資金供与による論文は塗りつぶした長方形、その他は中空の長方形で示してある。NEI の資金供与になる論文を引用する各特許(塗りつぶしの楕円形で示す)から当該論文まで線が引いてある。NEI の資金供与を受けていない論文への引用は表示していない。NEI の資金供与を受けた論文への引用と紛らわしくなるからである。グラフ上の引用の密度は NEI 資金供与研究の影響力を表現している。

35 研究の特許のうち 14 件は、NEI 助成金により安定的な支援を受けている研究者であるコロンビア大学の Laszlo Bito による合計 12 点の別々の論文を引用した。図 4-20B は Bito の論文にかぎってそれとの関連を示し、この研究代表者が技術に与える重要な影響を例示している。

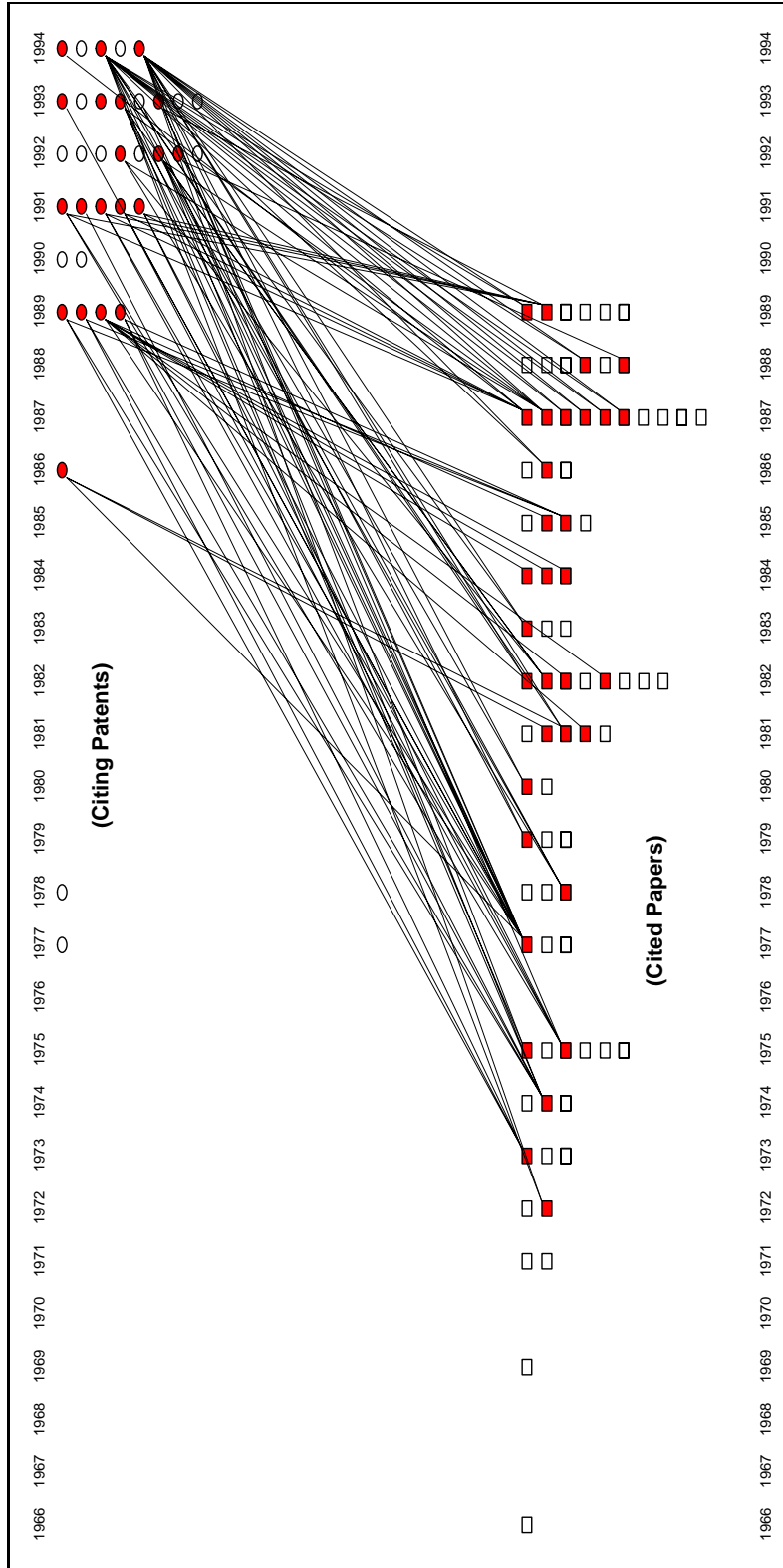
### コメント



ファルマシアとアブジョンは、ラタナプロストを1996年9月に医薬品ザラタンとして市場に出し980万ドルの売上を記録した。その後の四半期報告書では、1997年9月末の年間売上高は合計1億2,900万ドルに達したとしている。NEIの資金供与に直接帰せられる商業的影響の正確な部分はこれらのデータから決定するのは不可能だが、この簡潔な分析に示した証拠からも影響が大きいことがわかる。任意の個々の資金供与機関、科学者、発明家または会社などが経済的生産物の一部に貢献していることを追跡するには、広範で詳細な調査、例えば、基礎科学的探求から最終生産物のエンジニアリングやマーケティングにいたる当該生産物の一連の流れを調査する必要がある。しかし、本研究で講じた方法は何らかのアクターが与える重要な影響を明らかにする。

Figure 4-20A

Linkage Between Prostaglandin Glaucoma Technology and NEI-Funded Research

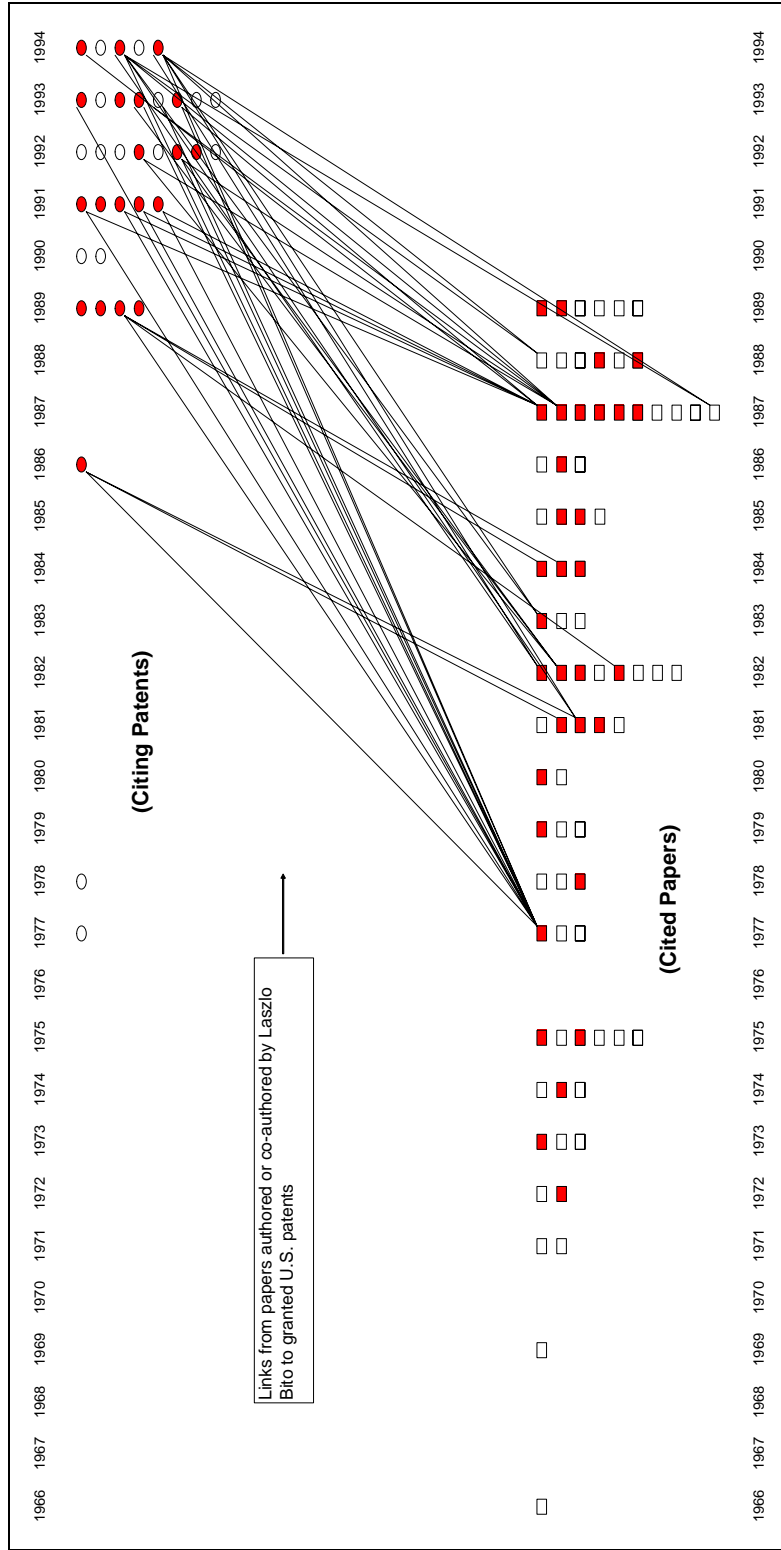


Legend:

- Patent citing NEI-funded paper
- Patent not citing NEI-funded paper
- Paper acknowledging NEI funding
- Paper not acknowledging NEI funding
- (with line) Patent citing paper
- (with line) Paper cited by patent

Notes: Papers are "References Cited-Other Publications" listed on the front page of the patent. Only papers that could be found in a library lookup are included. Among the papers found and represented above, 31 gave no indication of a funding support source. Excluded from this figure are references to patents, to books, and to publications that could not be found in a library search.

Bito's Role in Science Literature and Technology of Prostaglandins in Treatment of Glaucoma



Patent citing paper supported by NEI funding  
 Patent not citing paper supported by NEI funding  
 Paper acknowledging NEI funding  
 Paper not acknowledging NEI funding

Notes: Papers are "References Cited-Other Publications" listed on the front page of the patent. Only papers that could be found in a library lookup are included. Among the papers found and represented above, 31 gave no indication of a funding support source. Excluded from this figure are references to patents, to books, and to publications that could not be found in a library search.

## 10. 科学的基盤を NEI 資金供与研究にたどれる派生的技術の経済的影響の検証

報告者: Peter Kroll & Francis Narin, December 1997

### 目的

NEI 研究所の資金供与研究が眼科医療に直接関係しない技術に及ぼす影響の評価。

### 方法論

初期の調査で作成された眼科医療技術特許により引用された NEI 資金供与論文のセットを利用することにより、1983 年～96 年以降に発行された特許のうちそれらの論文を引用する 525 件 (すでに特定されている眼科医療特許 390 件を含む) すべてが特定された。さらに、米国で書かれた論文で、1987 年～88 年と 1993 年～94 年に発行された特許が引用している論文をすべて探索し、このうち資金源が NEI でない論文を特定した。また、眼科医療技術特許が引用していない他の NEI 資金供与論文も上記の特許期間について決定された。セット中の重複を考慮に入れ、かつこれらの組合せの中で新規に特定された多数の眼科医療特許を除いて、非眼科医療関連論文 262 件の最終セットを特定し、これを検証の対象にした。

### 結果

これらの組合せにおいて NEI 文献を引用する 679 件の論文のうち、39%は眼科医療関連以外のものであった。この「派生」特許のうち、NEI 論文を引用している特許譲受者の上位3位は大学で、4位から 13 位までは法人であった。

派生特許は主として他の生物医学分野に含まれていた。それらはさらに技術および応用分野に区分されていた(図 4-21 参照のこと)。医薬品や基礎科学/研究が技術区分の多くを占めていた。診断学および医療/外科技術も大きな分野であった。「その他」の区分に分類されたのは6件の論文にすぎなかった(水脈/鉍脈探知器および高度デジタルテレビシステムなど)。

特許を特定の生物医学応用分野別に分類すると、当該特許の半数以上は生物医学分野、例えば、医療応用別には分類できない研究技術などの最も基礎的な分野であった。ガンに応用で

きる特許は合計の約5分の1を占め、傷/皮膚病、心臓/循環器、および伝染病の分類がそれぞれ特許の10%から12%を占めている。

これらの派生特許のいくつかは他の特許による引用度が高く、潜在的には重要な技術である。

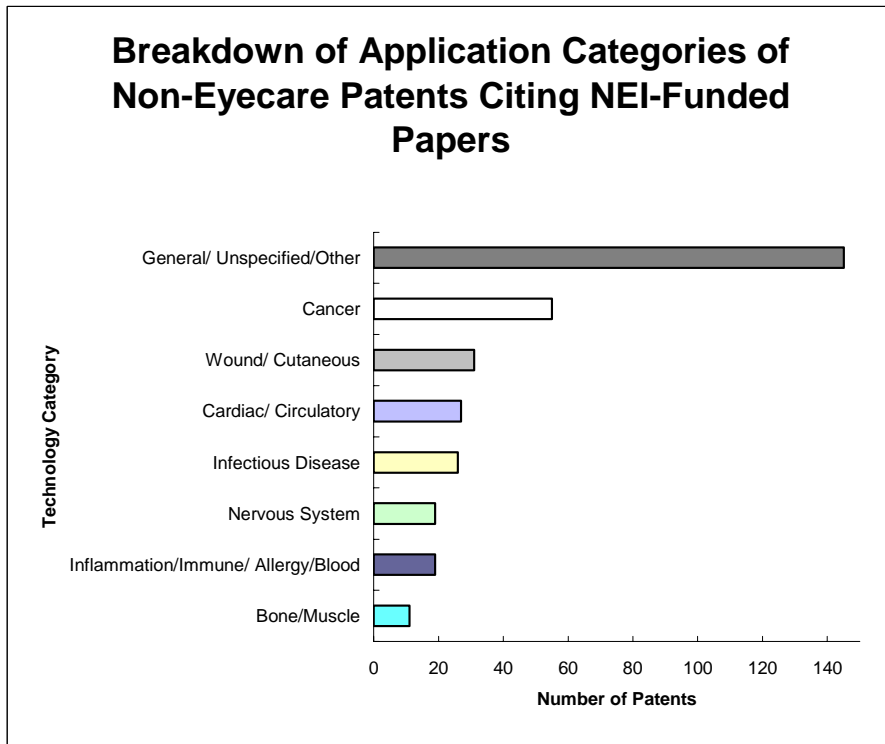
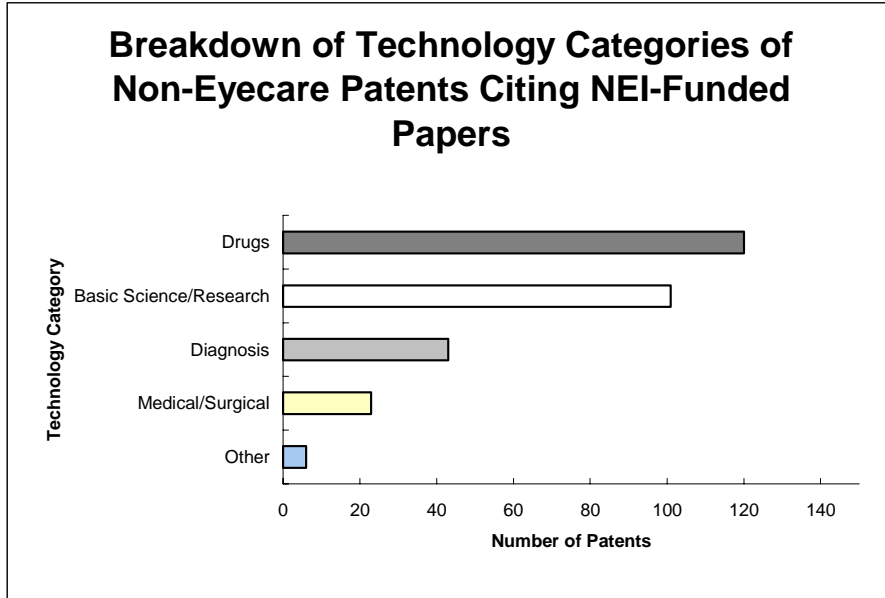
#### コメント

これらの派生特許の経済的影響の分析は、本研究の範囲を越えるが、これにより眼科医療技術以外の NEI 資金供与研究の影響の定量的尺度を検討することができるだろう。これらは、眼科医療以外の医療関連特許に対する大きな影響とともに単一機関の研究資金供与の広範な広がりを示している。

研究資金源特定の性格は利用可能データの組合せに左右される。とくに、この方法には1987年～88年、1993年～94年の年を除く年における特許、すなわち、それ以前には眼科医療特許が引用してきたものと特定されなかった論文を引用する特許が含まれていない。含まれていない特許の別の事例としては、NEIの資金供与によるものであり、かつ米国以外の著者の機関の論文を引用する特許があげられよう。

图 4-21

Distribution by Categories



## 11. STTR 技術に対する公的資金供与研究の影響の検証

**Peter Kroll, Grace Ault, and Francis Narin, "Tracing the Influence of Basic Scientific Research on biotechnology Patents: A Case Study of Signal Transduction and Transcriptional Regulation (STTR)," Patent World, issue 100, March 1998, 38-46.**

### 目的

本事例研究は重要技術の基礎になっている科学を特定するために、それら技術や、さらには研究に資金を供与する政府機関の立場から振り返る一連の事例研究の一つであった。ここで検証する研究はシグナル伝達と転写制限(STTR)分野、すなわち、新規医療開発のための最も有望な分野の1つにおける研究を検討した。がん、心臓血管病および関節炎やぜんそくなどの炎症の病気はすべて細胞の成長または活動を管理する情報伝達経路への信号発信の異常に関連している。STTR 特許が引用する研究の起源、とくに資金供与源は非常に興味深い。

### 方法論

1991-96年のSTTR特許200件がキーワードや会社名検索で特定された。このサンプルは引用する特許以外の参照が多いため、1995年の特許に限定した。62の特許にはそれぞれ平均26.2件の特許以外の参照(参照総計1,627点)があった。さらなる検証の結果、これらのうち1,457件は科学ジャーナルへの参照で、1,361件の別々の論文が得られた(参照の中には同じ論文を引用したものもあった)。

### 結果

特許の所有権は譲受者のあいだに広く拡散し、62件の特許を所有する組織(共同譲受者も含まれる)は54もある。譲受者の半数以上は、民間の営利法人で、このことからこの分野では商業開発に大きな関心が集まっていることがわかる。著者の機関には大学や研究機関のほか、医薬品会社数社を含んでいた。NIHの研究所は著者の機関として92回、ハーバード大学医学部は90回言及された。当該分野の論文は影響度が大きく、基礎指向が強いジャーナルに掲載された。

ここでは著名なジャーナルが上位を占めている。資金供与源は NIH が支配的であった。資金供与組織のうち上位3位、上位6位のうち5、および上位 11 位のうち8を占めたのは NIH の研究所であった。図 4-22A は、図書館にあった論文合計 1,297 件のうち 49%に NIH 資金供与への謝辞が含まれていることを示している。NIH の資金供与対象にはなっていないが、全体の 5%にあたる NIH 内科学者によって作成された論文も NIH 支援の対象として挙げられよう。

統計の別の見方は引用特許の観点である(図 4-22B 参照)。ジャーナル参照を含む 60 件の特許のうち、54 件は NIH 資金供与論文または NIH の著作論文を少なくとも1点引用していた。最後に NIH と NIH 以外の機関の論文とのあいだの指標尺度を比べてみると、NIH 支援論文は、より基礎的で影響力のある下位分野で発表され、また下位分野の中で比べても NIH 以外の研究所より基礎的で、影響力のある論文であった。

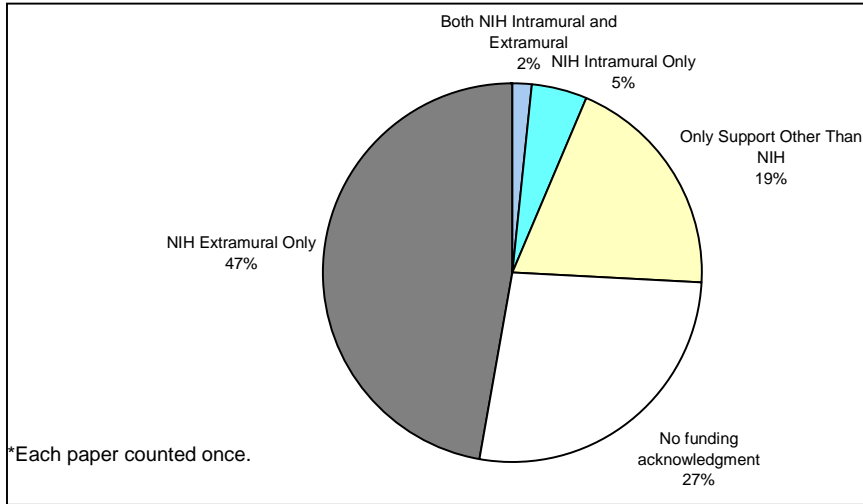
#### コメント

STTR の分野は成熟しておらず、商業製品の経済的影響評価の対象にならなかった。この傾向は当該分野における官民の特許所有権のバランスと一致している。NIH の中心的役割は明らかになった。



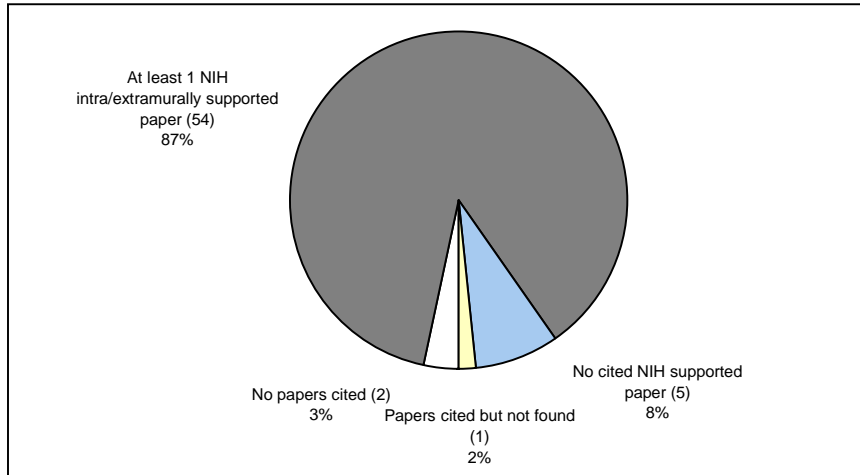
☒ 4-22A

NIH Funding Support for Papers Cited in 1995 STTR Patents: Counted by Paper



☒ 4-22B

NIH Funding Support for Papers Cited in 1995 STTR Patents: Counted by Patent



## 12. 特許に引用された論文に対するオーストラリア政府の資金供与の特徴の検証

報告者: Francis Narin, Michael Albert, Peter Kroll & Diana Hicks, February 2000

### 目的

本研究ではオーストラリアの特許取得制度におけるパタンおよびオーストラリア特許と公的支援対象科学研究とのあいだの関連性を検証した。ここに紹介した報告の焦点はオーストラリア政府(オーストラリア研究評議会および CSIRO)の支援研究に関する課題である。

### 方法論

当該研究は次の3つの問題で構成されている。オーストラリアの特許と世界の科学研究とのあいだの関連性、世界の特許とオーストラリアの科学研究との間の関連性、およびオーストラリア特許の長所と特徴。すべての場合において特許は米国特許・商標局が発行した特許に限定された。CHI リサーチ社の特許・論文リンケージ・データベースが上記の問題の分析のために使用された。NASA と Battelle 研究所の2つの米国組織が CSIRO 特許指標の比較対象とされた。

### 結果

オーストラリア政府はオーストラリアにおける特許取得制度を管理している。CSIRO は同国のトップ特許取得組織としての地位を占め、また CSIRO 以外の同国政府の特許をしのいでいる。オーストラリア特許では部門(セクター)間引用はほとんどない(図 4-23A 参照のこと)。同国の特許はほとんど引用されていない。

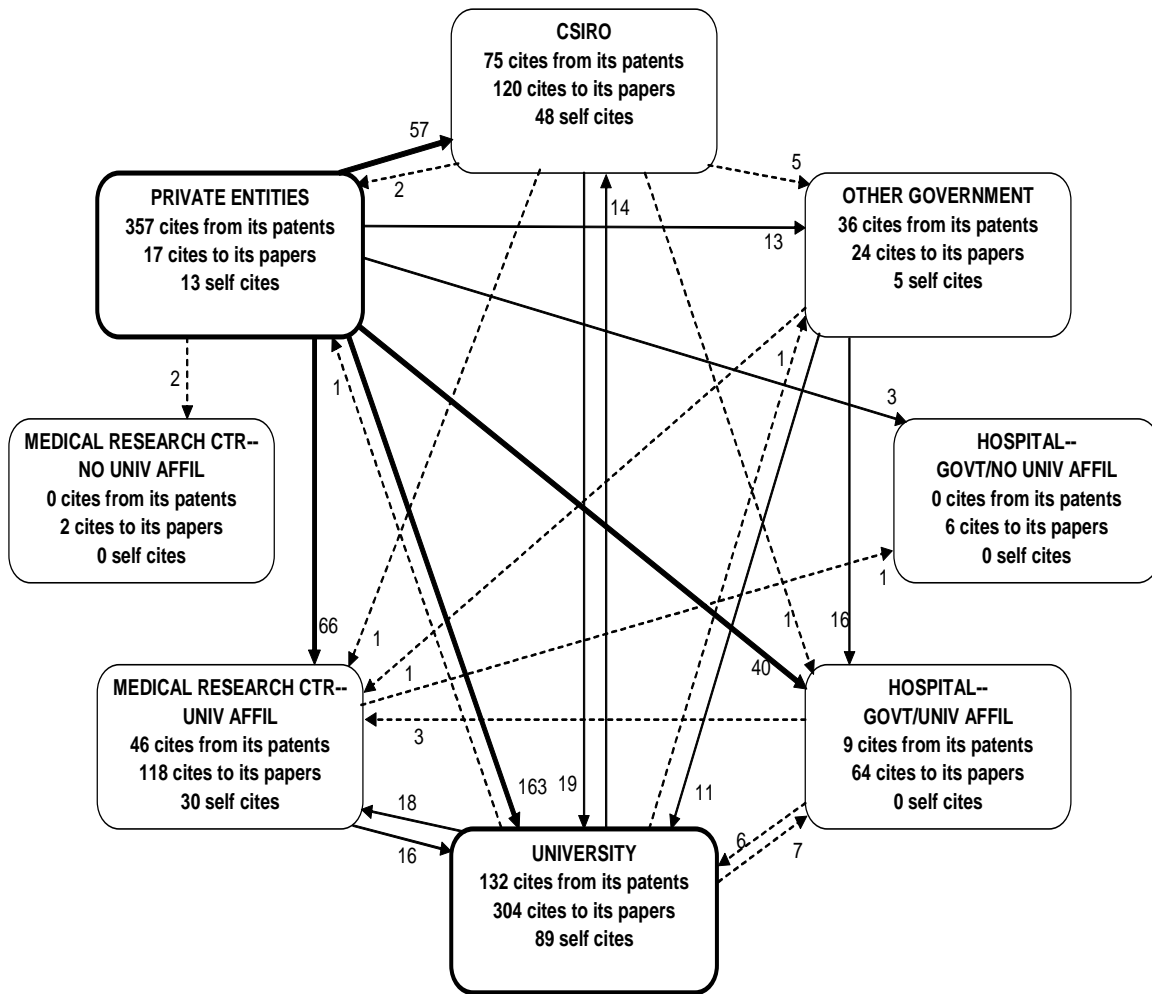
同国科学研究の特許による引用については、同国の科学への引用全体の約 95%が同国の公的科学研究に対するものであり、63%は大学およびそれらの関連医療センター、14%は CSIRO、5%は民間会社そして残りはその他さまざまな公的部門の研究機関に対するものである。あらゆる研究分野にわたって、CSIRO とオーストラリアの大学からの論文は、特許に引用された同国の科学の中でも際だっている。オーストラリア連邦(CSIRO を含む)は同国の論文に引用される同国特

許(少なくとも1名のオーストラリア人発明家を含む)において2位を占めている。資金供与組織、明確な資金供与組織の記載がない論文については著者の機関の一覧表からは、とくに同国政府の機関および公的機関一般が特許に引用された同国科学研究の最大支援機関であることがわかる。図 4-23A に見られる公的部門の論文への高い特許引用度と図 4-23B に見られる特許間引用の低さを比べてみるとわかる。

#### コメント

本研究はオーストラリアにおける公共政策の議論に著しい影響を及ぼした。本研究が公表された時期がイノベーションをめぐる国民的議論が高まった時期と一致していたからである。この点については次のウェブを参照のこと。 [http://www.arc.gov.au/publications/arc\\_pubs/00\\_02.pdf](http://www.arc.gov.au/publications/arc_pubs/00_02.pdf)

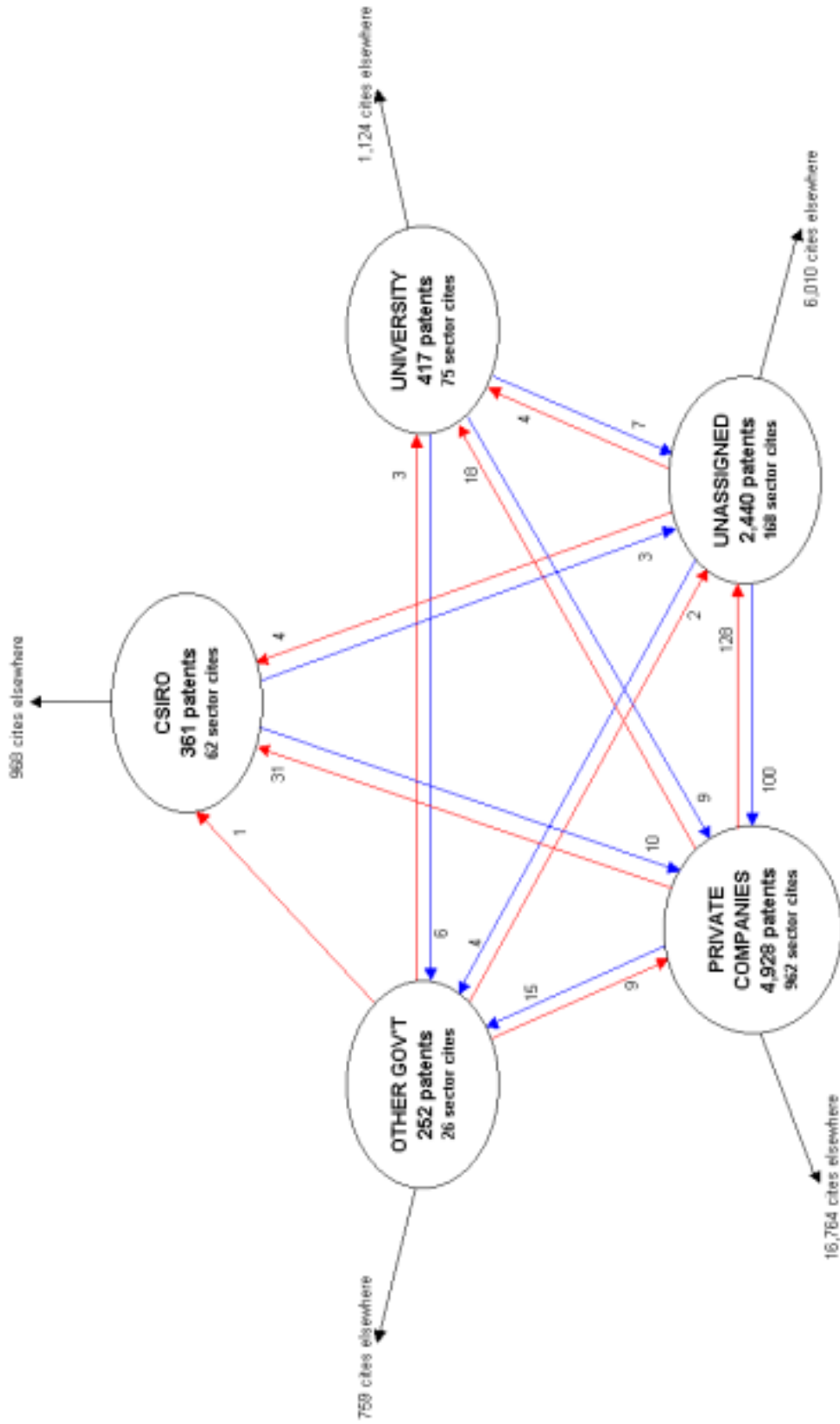
Citation of Papers by Patents Between Australian Sectors



Notes: Counts are of references by patents to papers.  
 Patents are counted fractionally; counts less than 0.5 have been omitted.  
 Patents by private entity inventors include foreign assignees.  
 Counts may not add up due to rounding

4-23B

Cross-sector Citations Among Invented 1979-1998 U.S. Patents



## 第5章 参考文献

(太字で示したものは、附属資料で再掲している)

Advisory Board for the Research Councils, The Royal Society, Economic and Social Research Council.  
*Evaluation of National Performance in Basic Research*, Department of Education and Science, London, 1986.

Anderson, Richard C., Francis Narin, and Paul R. McAllister. "Publication Ratings vs. Peer Ratings of Universities," *Journal of the American Society for Information Science*, 29, 91-103, 1978. Reprinted in *Key Papers in Information Science*, Belver C. Griffith, ed., Knowledge Industry Publications, Inc. 1980.

Averch, H.A. "Measuring the Cost-Efficiency of Basic Research Investment: Input-Output Approaches," *Journal of Policy Analysis and Management*, 6, 3, 342-361, 1987.

Averch, H.A. "Exploring the Cost-Efficiency of Basic Research Funding in Chemistry," *Research Policy*, 18, 165-172, 1989.

Averch, H. "Annotated Bibliography on Evaluation of Research, 1985-1990," in Bozeman & Melkers eds., 279-300, 1993.

Averch, H. "Criteria for Evaluating Research Projects and Portfolios," in Bozeman & Melkers, eds., 263-277, 1993.

Bozeman, B. & Julia Melkers, eds. *Evaluating R&D Impacts: Methods and Practice*, Kluwer, Dordrecht, 1993.

Buderi, Robert, John Carey, Neil Gross and Karen Lowry Miller. "Global Innovation: Who's in the Lead?" *Business Week Patent Scoreboard*, August 3, 1992.

Carpenter, Mark and Francis Narin. *Utilization of Scientific Literature by U.S. Patents*, final report to National Science Foundation on contract no. PRM-7801694, 1978.

Carpenter, Mark, M Cooper and Francis Narin. "Linkage Between Basic Research Literature and Patents." *Research Management*, 23, 30, 1980.

Carpenter, Mark P., Francis Narin and Patricia Woolf. "Citation Rates to Technologically Important Patents," *World Patent Information*, 4, 160-163, 1981.

Carpenter, Mark & Francis Narin. "Validation study: Patent Citations as Indicators of Science and Foreign Dependence," *World Patent Information*, 5, 180, 1983.

- Carpenter, Mark. "Patent Citation as Indicators of Scientific and Technological Linkages." AAAS Annual Meeting. Detroit, MI, USA, May 30, 1983.
- Collins, Peter & Suzanne Wyatt. "Citations in Patents to the Basic Research Literature" *Research Policy*, 17, 65, 1988.
- Committee on Science, Engineering, and Public Policy (COSEPUP). *Evaluating Federal Research Programs: Research and the Government Performance and Results Act*, National Academy Press: Washington, D.C., 1999.
- Coy, Peter, and John Carey. "The Global Patent Race Picks Up Speed," *Business Week Patent Scoreboard*, August 16, 1993.
- Cozzens, S.E., Popper, S., Bonomo, J., Koizumi, K., Flanagan, A. *Methods for Evaluating Fundamental Science*. RAND/CTI DRU-875/2-CTI, Washington, DC, 1994.
- Cozzens, S.E. *Assessment of Fundamental Science Programs in the Context of the Government Performance and Results Act GPRA*, Rand, Washington D.C., MR-707.0-OSTP, October 1995a.
- Cozzens, S.E. "U.S. Research Assessment: Recent Developments", *Scientometrics*, 34 3, 351-362, 1995b.
- Deng, Zhen, Baruch Lev, and Francis Narin. "Science & Technology Indicators as Predictors of Stock Performance," *Financial Analysts Journal*, 55, 3, 20-32, May/June 1999.
- Department of Defense. "Project Hindsight", Office of the Director of Defense Research and Engineering, DTIC Report No. AD495905, October 1969. See also, Sherwin, C.W. and R.S. Isenson, "Project hindsight: defense department study of the utility of research," *Science*, 156, 1571-1577, 1967.
- Dietz, James S., Ivan Chompalov, Barry Bozeman, Eliesh O'Neil Lane, and Jongwon Park. "Using the Curriculum Vita to Study the Career Paths of Scientists and Engineers: An Exploratory Assessment," *Scientometrics*, 49, 3, 419-442, 2000.
- Ellis, P., G. Hepburn and C. Openheim. "Studies on Patent Citation Networks," *Journal of Documentation*, 34, 1, 12-20, March 1978.
- Garfield, Eugene. "Citation Indexes for Science," *Science*, 122, 108-111, 1955.
- Geisler, E. *The Metrics of Science and Technology*, Quorum Books, Westport CN, 2000.
- Georghiou, L. and D. Roessner. "Evaluating Technology Programs: Tools and Methods," *Research Policy*, 29, 657-678, 2000.

Glanzel, W., A. Schubert and H.-J. Czerwon. "An Item-by-Item Subject Classification of Papers Published in Multidisciplinary and General Journals Using Reference Analysis," *Scientometrics*, 44, 3, 427-440, 1999.

Griliches, Zvi. "Market Value, R&D and Patents". *Economics Letters*, 7, 183-187, 1981.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg. *Market Value and Patent Citations: A First Look*, Paper prepared for the Conference on Intangibles and Capital Markets, New York University, 1998.

Harhoff, Dietmar, Francis Narin, F. M. Scherer, and Katrin Vopel. "Citation Frequency and the Value of Patented Inventions," *The Review of Economics & Statistics*, 81, 3, 511-515, August 1999.

**Hicks, D. "The Difficulty of Achieving Full Coverage of International Social Science Literature and the Bibliometric Consequences," *Scientometrics*, 44, 2, 193-215, 1999.**

**Hicks, D., A. Breitzman, K. Hamilton, F. Narin. "Research excellence and patented innovation," *Science and Public Policy*, 275, 310-320, 2000.**

Hodgdon, J.D., S.J. Fitzsimmons, L.C. Kerpelman, M.C. Jerrett, D. Deal, J. Weinberg. *Methods for the Strategic Evaluation of Research Programs: The State-of-the-Art – Annotated Bibliography*, report to National Science Foundation on contract No. PRA 8400688 Abt Associates Inc.: Cambridge MA, 1985.

Inhaber H. and K. Przednowek. "Quality of Research and the Nobel Prizes," *Social Studies of Science*, 6, 33-50, 1976.

Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson. "Geographic Localization of Knowledge Spillovers and Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108, 3, August 1993.

Katz, J.S. and Diana Hicks. "How much is a collaboration worth? A calibrated bibliometric model," *Scientometrics*, 40, 541-554, 1997.

Kerpelman, L.C. & S.J. Fitzsimmons. *Methods for the Strategic Evaluation of Research Programs: The State-of-the-Art – Analytical Overview Report*, report to National Science Foundation on contract No. PRA 8400688 Abt Associates Inc.: Cambridge MA, 1985.

Kleinberg, J.M. "Hubs, Authorities and Communities", *ACM Computing Surveys*, 31, 4, December, 1999, available at: <http://dev.acm.org/pubs/articles/journals/surveys/1999-31-4es/a5-kleinberg/a5-kleinberg.pdf>.

Kostoff, R. "Evaluation of Proposed and Existing Accelerated Research Programs by the Office of Naval Research," *IEEE Transactions on Engineering Management*, 35, 4, 1988.

Kostoff, R. "Evaluating Federal R&D in the United States", in Bozeman & Melkers, eds. 163-178, 1993.



Kostoff, R.N. "Federal Research Impact Assessment: State-of-the-Art," *Journal of the American Society for Information Science*, 456, 428-440, 1994.

Link, A.N. *Evaluating Public Sector Research and Development*, Praeger, Westport, CN, 1996.

Logsdon, J.M. & C.B. Rubin. "Research Evaluation Activities of Ten Federal Agencies," *Evaluation and Program Planning*, 11, 1-11, 1988.

Martin, B.R. and J. Irvine. "Assessing Basic Research: Some Partial Indicators of Scientific Progress in Radio Astronomy," *Research Policy*, 12, 61-90, 1983.

**McAllister, Paul R., Francis Narin and James G. Corrigan. "Programmatic Evaluation and Comparison Based on Standardized Citation Scores," *IEEE Transactions on Engineering Management*, 30, 4, 205-211, November, 1983.**

Melkers, J. *Bibliometrics as a Tool for Analysis of R&D Impacts*, in Bozeman & Melkers, eds. 42-61, 1993.

Melkers, J. and D. Roessner. "Politics and the Political Setting as an Influence on Evaluation Activities: National Research and Technology Policy Programs in the United States and Canada," *Evaluation and Program Planning*, 20 1, 57-75, 1997.

Meyer, Martin. "Does Science Push Technology? Patents Citing Scientific Literature" *Research Policy*, 29, 409-434, 2000.

Narin, Francis. *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity*, Contract NSF C-627, National Science Foundation. March 31. Monograph: 456 NTIS Accession #PB252339/AS, 1976.

Narin, Francis and Elliot Noma. "Is Technology Becoming Science," *Scientometrics*, 7, 3, 369-381, 1985.

Narin, Francis, Elliot Noma and Ross Perry. "Patents as Indicators of Corporate Technological Strength," *Research Policy*, 16, 143-155, 1987.

Narin, Francis. Globalization of Research, Scholarly Information and Patents – Ten Year Trends. In: Proceedings of the North American Serials Interest Group NASIG 6<sup>th</sup> Annual Conference June 14-17, *The Serials Librarian*, 21, 2-3, 1991.

Narin, Francis & Dominic Olivastro. "Status report: Linkage Between Technology and Science," *Research Policy*, 21, 237, 1992.

**Narin, F. and K.S. Hamilton. "Bibliometric Performance Measures," *Scientometrics*, 36, 3, 293-310, 1996.**

**Narin, Francis, Kimberly S Hamilton and Dominic Olivastro. "The Increasing Linkage between U.S. Technology and Public Science," *Research Policy*, 263, 317-330, 1997.** Reprinted in the *AAAS Science and Technology Yearbook*, 1998. **"Study Finds Public Science is Pillar of Industry", *New York Times Science Times Section*, Tuesday, May 13, 1997.**

Narin, Francis and Dominic Olivastro. "Linkage Between Patents and Papers: an Interim EPO/US Comparison," Prepared for *Proceedings of the Sixth Conference of the International Society for Scientometrics and Infometrics*, June 16-19, 1997, Jerusalem, Israel. Published in a special edition of *Scientometrics*, 41, 1, 51-59, 1998.

National Science Foundation. *Science Indicators 1972*, Report of the National Science Board, 1973, and subsequent biennial *Science and Technology Indicators* reports.

Office of Technology Assessment. *Research Funding as an Investment: Can We Measure the Returns? – A Technical Memorandum*, Washington, DC: U.S. Congress, Office of Technology Assessment, OTA-TM-SET-36, April 1986.

Office of Technology Assessment. *Federally Funded Research: Decisions for a Decade*, OTA-SET-490 Washington, DC: U.S. Government Printing Office, May 1991.

Patent and Trademark Office, U.S. Department of Commerce. *Technology Assessment and Forecast*, Sixth Report, June 1976.

Pinski, Gabriel, and Francis Narin. "Citation Influence for Journal Aggregates of Scientific Publications: Theory, with Application to the Literature of Physics," *Information Processing and Management*, 12 5 97-312, 1976.

Popper, S. W. *Economic Approaches to Measuring the Performance and Benefits of Fundamental Science*, Rand, Washington D.C., MR-708.0-OSTP, October 1995.

Reisner, P. A Machine Stored Citation Index to Patent Literature Experimentation and Planning. In: *Proceedings of Automation and Scientific Communications Annual Meeting 1963*, H.P. Lunh ed.. American Documentation Institute, Washington, DC, 1965.

Research Value Mapping Project. *Qualitative-Quantitative Case Studies of Research Projects Funded by the Office of Basic Energy Sciences*, final report to the Office of basic Energy Sciences, Department of Energy, February 19, 1999.

Roessner, D. *Use of Quantitative Methods to Support Research Decisions in Business and Government*, in Bozeman & Melkers, eds. 179-205, 1993.

Roessner, D., B. Bozeman, I. Feller, C. Hill, N. Newman. *The Role of NSF's Support of Engineering in Enabling Technological Innovation*, first year final report for National Science Foundation SRI International: Arlington VA, January 1997.

Roessner, D., C. Ailes, I. Feller, L. Parker. "How Industry Benefits from NSF's Engineering Research Centers", *Research, Technology, Management*, 41, 5, 40-44, 1998a.

Roessner, D., R. Carr, I. Feller, M. McGeary, N. Newman. *The Role of NSF's Support of Engineering in Enabling Technological Innovation: Phase II*, final report to National Science Foundation SRI International: Arlington VA, May 1998b.

Roessner, David. "Quantitative and Qualitative Methods and Measures in the Evaluation of Research," *Research Evaluation*, 8, 2, August 2000.

Rogers, J.D. & B. Bozeman. " 'Knowledge Value Alliances': An Alternative to the R&D Project Focus in Evaluation," *Science, Technology and Human Values*, 26, 1, 2001.

Patent and Trademark Office, U.S. Department of Commerce. *Manual of Patent Examining Procedures*, Section 904.02, 6<sup>th</sup> edition, 1995.

Seglen, P.O. "Why the Impact Factor of Journals Should not be Used for Evaluating Research," *British Medical Journal*, 314, 498-502, 1997.

Trajtenberg, Manual. "A Penny for your Quotes: Patent Citations and the Value of Innovations," *Rand Journal of Economics*, 21, 11, 1990.

## 附属資料 主要文献の概要

Abt Associates.

A list of their evaluations with short summaries can be found at:

<http://www.abtassociates.com/areas/science-projects.html>. This consulting firm has conducted a large number of evaluations for federal agencies.

Averch, H.A. "Measuring the Cost-Efficiency of Basic Research Investment: Input-Output Approaches," *Journal of Policy Analysis and Management*, 6, 3, 342-361, 1987.

Combines bibliometric and econometric approaches. In the author's words: Examines the strengths, weaknesses and assumptions of the Irvine-Martin approach to assessing the research productivity of institutions by counting their absolute numbers of citations and publications. Suggests that there would be payoff in looking at information return per dollar of agency investment. Carries out the first known example of an efficiency comparison for basic research in NSF's behavioural science program.

Averch, H.A. "Exploring the Cost-Efficiency of Basic Research Funding in Chemistry," *Research Policy*, 18, 165-172, 1989.

In the author's words: the first econometric study of research funding efficiency by a federal agency. Uses citations per dollar of research investment as the measure of efficiency. Shows that efficiency is determined by highly interactive variables such as investigators' experience and the labor intensity of a project. Shows there is enormous variation in efficiency and that at least for the given sample, some factors thought to be important by the scientific community such as quality of department are not.

Averch, H. *Annotated Bibliography on Evaluation of Research, 1985-1990*, in Bozeman & Melkers, eds. 279-300, 1993.

Comments in bibliography focus particularly on issue of use of evaluations. Averch takes the position that non-use stems from nature of policy making process. Thus improved methods are seen as rhetorical tools to sell studies, but do not increase the use of results in policy making. Averch is critical of the fact that studies are used to demonstrate that policy makers did a good job, but are not used to change resource allocations. Raises the issue of cost effectiveness of evaluation, given its non-use.

Averch, H. *Criteria for Evaluating Research Projects and Portfolios*, in Bozeman & Melkers, eds. 263-277, 1993.

Combines an economist's very rational way of looking at things with a perceptive view of the realities of life in government. Interesting is the proposal of practical strategies, or "algorithms" for evaluating three types of research: basic research, policy research for use in public decision making and innovation research. Each strategy is stripped to its core elements (p. 268).

Bozeman, B. & Julia Melkers eds. *Evaluating R&D Impacts: Methods and Practice*, Kluwer, Dordrecht, 1993.

Out of print and so not easy to obtain. Authors take positions on why evaluation methods are not more used. Implicit but ever present in these discussions is their limitation to the U.S. Federal context. OTA 1986 and Abt studies from mid-1980's are heavily relied upon.

Committee on Science, Engineering, and Public Policy (COSEPUP). *Evaluating Federal Research Programs: Research and the Government Performance and Results Act*, National Academy Press: Washington, D.C., 1999.

A discussion of GPRA from the leading advocate and contractor in the peer review area.

Cozzens, S.E. *Assessment of Fundamental Science Programs in the Context of the Government Performance and Results Act (GPRA)*, (Rand, Washington D.C., October 1995a, MR-707.0-OSTP).

Can be ordered from: <http://www.rand.org>. There is no charge.

Explains the requirements of the GPRA and how they relate to program evaluation. GPRA enshrined the distinction between outputs and outcomes and stated that outcomes are of primary interest. GPRA is problematic for research because the framework is that agencies set goals and measure their progress toward them. This makes sense for research activity (number of grants awarded, number of research students supported etc.), but not for the use and application of research which is largely out of an agency's control.

The most balanced and sophisticated treatment of evaluation methods in the GPRA context.

Cozzens, S.E. "U.S. Research Assessment: Recent Developments", *Scientometrics*, 34, 3, 351-362, 1995b.

This quick survey shows that quantitative evaluation studies are relatively rare in the United States. Most of the assessment of federal research programs is descriptive and expert judgement seems to be the favored method. GPRA requirements and the problems they have caused are discussed. Four agency pilot projects to develop acceptable indicators are described; the projects illustrate the underlying tension of agencies avoiding quantification and the Office of Management and Budget demanding measurable indicators.

Department of Defense. "Project Hindsight," Office of the Director of Defense Research and Engineering, DTIC Report No. AD495905, October 1969. See also, Sherwin, C.W. and R.S. Isenson, "Project Hindsight: Defense Department Study of the Utility of Research". *Science*, 156, 1571-1577, 1967.

Geisler, E. *The Metrics of Science and Technology*, Quorum Books, Westport CN, 2000.

There are chapters on: economic & financial, bibliometric, co-word, patents, peer review, and process outcomes (a variant of quantified case study), as well as chapters on application in industry, government, academia etc. Draws on a vast literature, covers a lot of ground and so wanders into general issues of science and technology policy. Takes an even handed, strengths and weaknesses approach.

Georghiou, L. and D. Roessner. "Evaluating Technology Programs: Tools and Methods," *Research Policy*, 29, 657-678, 2000.

The most recent review of evaluation methods reports newer developments, and acknowledges that the literature on R&D evaluation is now too vast to be comprehensively reviewed. Roessner is an influential U.S. practitioner and the review introduces work with which he has been associated.

Hodgdon, J.D., S.J. Fitzsimmons, L.C. Kerpelman, M.C. Jerrett, D. Deal, J. Weinberg. *Methods for the Strategic Evaluation of Research Programs: The State -of-the-Art - Annotated Bibliography*, report to National Science Foundation on contract No. PRA 8400688, Abt Associates: Cambridge MA, 1985.

Bibliography of articles on bibliometrics, economic or stochastic decision models, peer review, studies using several methods and others. Two to three paragraphs of commentary on each study.

Kerpelman, L.C. & S.J. Fitzsimmons. *Methods for the Strategic Evaluation of Research Programs: The State-of-the-Art - Analytical Overview Report*, report to National Science Foundation on contract No. PRA 8400688, Abt Associates: Cambridge MA, 1985.

A report focusing on how methods were used, how they were combined in use and how effective they were. Concludes:

- 1) peer review is the "touchstone". Bibliometrics has demonstrated utility but studies recommend it be used in conjunction with other techniques (the same is not said of peer review). Econometric techniques are frequently propounded, but save for cost benefit analysis, most techniques have not received widespread currency.
- 2) Evaluation techniques are combined in various ways.
- 3) Quantitative techniques appear to be used most often on applied research. Peer judgements are often successfully quantified.
- 4) Various outcomes, impacts and prestige measures are used.
- 5) Formal, strategic evaluation of research programs is not regularly conducted in government or industry.
- 6) Lots of methods are proposed, some are used, none are replicated.
- 7) The impact of evaluation on planning and resource allocation is limited because policy often overrides the conclusions of evaluations when decisions are made.

Kostoff, R. "Evaluation of Proposed and Existing Accelerated Research Programs by the Office of Naval Research," *IEEE Transactions on Engineering Management*, 35, 4, 1988.

Describes a method of evaluating a portfolio of agency projects using a peer review process. The process is standardized and fairly quantitative. Kostoff developed this process for the Department of Energy then transferred it to the ONR when he moved. The process is well regarded and was used more than once at DOE and ONR.

Kostoff, R. *Evaluating Federal R&D in the United States*, in Bozeman & Melkers, eds. 163-178, 1993.

A quick tour of main categories of method (peer review of grants, peer review of program outcomes, traces studies, bibliometrics, cost-benefit), with examples of peer review being the fullest, and the bib-

liometrics based on Narin's work. Nice overview on prevalence of techniques versus used, published studies. Lack of use discussed - relies on Averch.

Kostoff, R.N. "Federal Research Impact Assessment: State-of-the-Art," *Journal of the American Society for Information Science*, 45, 6, 428-440, 1994.

Material in this paper will also be found in Kostoff's *Handbook*.

Fairly standard. Author has some new critiques from his close-up contact with scientists in his work at the Navy. Contains an interesting calculation of the full costs of a peer review. Proposes that the future lies with large databases and co-occurrence techniques.

Kostoff, R.N. *The Handbook of Research Impact Assessment*, Seventh Edition, DTIC Report Number ADA296021, Office of Naval Research: Arlington, VA, 1997.

This and other work by Kostoff is at: <http://www.dtic.mil/dtic/kostoff/index.html>

The Handbook is the central work by Kostoff as most of his publications use material taken from this source. It is notable for its massive bibliography. It also contains extensive discussions of evaluations undertaken by Kostoff for the Navy, most of which seem to be peer review based. Under bibliometrics it contains extensive discussion of 1) Kostoff's work analyzing documents using word co-occurrences and 2) his work identifying direct and indirect impacts of research on other fields and application areas using a network analysis based on survey data.

Link, A.N. *Evaluating Public Sector Research and Development*, Praeger, Westport, CN, 1996.

Describes evaluations Link conducted for ATP. Chapter 2 covers evaluation issues: the evaluation method should be . . . Alternative economic methods - what they are and some of their problems - are discussed: multiple programming analysis, economic impact analysis and benefit-cost analysis (the later in the most depth). Exhibits a practitioner's sophisticated and detailed awareness of methodological problems.

Logsdon, J.M. & C.B. Rubin. "Research Evaluation Activities of Ten Federal Agencies," *Evaluation and Program Planning*, 11, 1-11, 1988.



Based on 44 interviews with research managers in the 10 major federal research funding agencies. Concludes that few federal agencies were making substantial use of research evaluation approaches other than informed peer judgement and review. An innovative and structured peer review is discussed as is the use of bibliometric approaches to research evaluation by NIH, and impact studies carried out by several agencies. The role of the National Academy of Sciences in promoting and undertaking peer review evaluations is clear.

McAllister, P.R., F. Narin & J.G. Corrigan. "Programmatic Evaluation and Comparison Based on Standardized Citation Scores", *IEEE Transactions on Engineering Management*, 30, 4, 205-211, 1983.

Derives a technique to create normalized citation scores comparable across fields using standard statistical methods.

Melkers, J. *Bibliometrics as a Tool for Analysis of R&D Impacts*, in Bozeman & Melkers, 42-61, 1993.

An even handed survey of bibliometric literature from someone without strong commitments for or against the method.

Melkers, J. and D. Roessner. "Politics and the Political Setting as an Influence on Evaluation Activities: National Research and Technology Policy Programs in the United States and Canada," *Evaluation and Program Planning*, 20, 1, 57-75, 1997.

A wide ranging review of research evaluation seen with the context of general program evaluation and with a comparative perspective that brings the U.S. situation into sharper focus. Explores how the fragmented U.S. political system creates a fragmented situation for program evaluation. Describes evaluation at the Office of Basic Energy Sciences (BES of DOE); NIST's Manufacturing Technology Centers, NSF's Engineering Research Centers, and NSF's Small Business Innovation Research program (SBIR).

Narin, F. and K.S. Hamilton. "Bibliometric Performance Measures," *Scientometrics*, 36, 3, 293-310, 1996.

The application of literature, patent and linkage bibliometrics to government research evaluation questions with reference to GPRA. Complements rather than repeats this report and is included in the appendix.

Office of Technology Assessment. *Research Funding as an Investment: Can We Measure the Returns?* - A Technical Memorandum, Washington, DC: U.S. Congress, Office of Technology Assessment, OTA-TM-SET-36, April 1986.

Available at: [http://www.wws.princeton.edu/~ota/disk2/1986/8622\\_n.html](http://www.wws.princeton.edu/~ota/disk2/1986/8622_n.html)

A classic. Perhaps the best review of economic and bibliometric measures and their use in the United States. Discusses the results of many studies, reaches measured, balanced and accurate conclusions and places the discussion in the broader policy context.

Office of Technology Assessment. *Federally Funded Research: Decisions for a Decade*, OTA-SET-490, Washington, DC: U.S. Government Printing Office, May 1991.

Available at: <http://www.wws.princeton.edu/~ota/>

Contains a five page discussion of evaluation which seems to rely on Chubin and Averch as contractors. Most of the discussion concerns practice in other countries.

Popper, S. W. *Economic Approaches to Measuring the Performance and Benefits of Fundamental Science*, Rand, Washington D.C., MR-708.0-OSTP, October 1995.

Can be ordered from: <http://www.rand.org>. There is no charge.

Discusses methods to measure the economic and social returns from investment in fundamental research. Bibliometrics are addressed, but of course bibliometrics does not measure economic returns. Report is a very good introduction to economic approaches. It offers a sophisticated, in depth discussion of several studies with careful consideration of strengths and weaknesses.

Research Value Mapping Project. *Qualitative-Quantitative Case Studies of Research Projects Funded by the Office of Basic Energy Sciences*, final report to the Office of Basic Energy Sciences, Department of Energy, February 19, 1999.

This and related papers not listed here can be obtained from:

<http://rvm.pp.gatech.edu/papers/index.html>. Juan & Rogers, 2000 is related.

An evaluation based on 28 case studies. The evaluation was quantitative in that answers to standardized questions were quantified and statistical methods were used to find patterns of program character-

istics associated with success. Four lessons were learned about the BES program. 1) Stable BES funding is invaluable to the field. 2) New organizational designs contribute to project effectiveness. 3) Basic research often spills over, resulting in new directions. 4) Interdisciplinary work requires a different management concept. 5) University-based projects differ little from those at federal labs.

Roessner, D. *Use of Quantitative Methods to Support Research Decisions in Business and Government*, in Bozeman & Melkers, 179-205, 1993.

One paragraph descriptions of studies undertaken for US federal agencies. Attributes lack of use to weakness in methods.

Roessner, D., B. Bozeman, I. Feller, C. Hill, N. Newman. *The Role of NSF's Support of Engineering in Enabling Technological Innovation*, first year final report for National Science Foundation, SRI International: Arlington VA, January 1997.

Available at: <http://www.sri.com/policy/stp/techin/>

Roessner, D., C. Ailes, I. Feller, L. Parker. "How Industry Benefits from NSF's Engineering Research Centers", *Research, Technology, Management*, 41, 5, 40-44, 1998a.

Report of a survey of companies involved with university research centers. Assesses their expectations against the benefits they gained.

Roessner, D., R. Carr, I. Feller, M. McGearry, N. Newman. *The Role of NSF's Support of Engineering in Enabling Technological Innovation: Phase II*, final report to National Science Foundation, SRI International: Arlington VA, May 1998b.

Available at: <http://www.sri.com/policy/stp/techin2/>

Roessner, David. "Quantitative and qualitative methods and measures in the evaluation of research," *Research Evaluation*, 8, 2, August 2000.

In the author's opinion, quantitative indicators pose dangers. Examples concerning evaluation of programs seeking to encourage collaboration are given.

Rogers, J.D. & B. Bozeman. "'Knowledge Value Alliances': An Alternative to the R&D Project Focus in Evaluation," *Science, Technology and Human Values*, 26, 1, 2001.

There are more related papers than are listed here. See Research Value Mapping, 1999 for URL to obtain related papers.

A presentation of the rationale behind the Research Value Mapping and a development of theory.

Rogers & Bozeman propose a new approach to research evaluation. The ideas arose out of an evaluation for the Department of Energy. He argues that projects and programs are most often bureaucratic fictions that bear little relation to the nature of research work. He concludes that evaluation should move closer to research as it is conducted by taking as its unit of analysis the alliances of researchers and others that work on a research topic and its application. The article discusses these ideas and several examples of networks of different types. The approach is well grounded in the most advanced sociology of science and addresses the complex intermingling of fundamental and application activities and institutions in research today. This complexity has come to forefront in the decades since research evaluation methods were first developed. If further work can establish the characteristics of alliances that are associated with economic and social benefits, the ideas may serve to move evaluators into a more realistic treatment of the nature of research today.



# 用語解説



➤ ATP (Advanced Technology Program: 先端技術プログラム)

米国政府の民間企業に対する産業技術支援プログラム。ATP は商務省の NIST (National Institute of Standards and Technology: 国立標準・技術研究所) が所轄している。

ATP は、民間と連邦政府機関が共同で行うハイリスクな研究開発に資金(いわゆるシードマネー)を提供する。ATP の対象はあくまでも基礎技術の開発であり、製品開発は助成対象にはならない。ただし民間主導で官民パートナーシップを推進するプログラムであり、営利企業が主体となってプロジェクトの企画・提案・実施を進め、民間負担分の資金を拠出する。それによって、革新的な製品、サービス、あるいは産業プロセスの開発を推進し、米国の企業および産業の機会拡大と競争力強化を図ることが狙いとされている。

なお、ATP は 1988 年の包括通商競争力法 (Omnibus Trade and Competitiveness Act of 1998) によって創設され、1991 年の米国技術卓越法 (American Technology Preeminence Act of 1991) に基づいて修正された。

米国の産業政策は市場メカニズムへの介入を避けてきたため、ATP は議会をはじめとする様々な場で批判や攻撃的となってきたが、それゆえに積極的に評価プログラムの開発に取り組んできたと言うことができよう。

➤ 計量文献学 (Bibliometrics: ビブリオメトリクス)

広義には文献を様々な側面から定量的に分析する研究領域を指す。文献計量学あるいは計量書誌学、書誌計量学とも呼ばれる。日本語の「計量文献学」という語は、文学作品や古文献などの文体や用語の統計的な研究を指す場合もあるが、本報告書がとりあげたような科学技術文献の統計的研究を指す「計量文献学」の語は、図書館情報学に起源を持つ “Bibliometrics” の訳語である。

科学技術文献の書誌情報には、掲載誌名、掲載時期、著者名やその所属機関名、その所在地、引用文献などが含まれ、科学技術活動についての社会科学的な分析研究の材料として重要である。そのため、科学技術文献を対象とする計量文献学では、個別文献のテキスト内容の分析よりも、多数の文献の書誌情報についての分析が主流となっている。特に、引用分析 (citation analysis) は、科学技術の特徴である知識の相互依存性、研究者のコミュニケーション、などの定量的分析を可能にする。また、論文の被引用回数は、論文の与えた影響度を示す指標と考えられ、さらにはそれによって研究の質を評価しようとする試みは広く



行われている。

このような科学技術文献の定量的分析研究は、Scientometrics (科学計量学ないし計量科学論)と呼ばれることも多くなっている。Scientometrics という語は、科学技術に関するあらゆる定量的分析研究の呼称となり得るが、実際には計量文献学的研究を指す場合が多い。このことは、計量文献学が、科学技術に関する定量的分析研究の中核に位置づけられていることを示している。さらには、科学技術文献を対象とした研究を中心に発展してきた計量文献学の諸手法は、最近では、特許の書誌的情報やインターネット上の情報を対象とする分析にも適用されている。

計量文献学は研究評価における情報提供のための有効なツールでもあり、国レベルの研究開発水準から、研究開発プログラムやプロジェクトの成果とそのインパクトの同定、あるいは、機関やグループの研究開発活動の変遷や特徴といった様々なレベルでの分析に用いられている。計量文献学は欧米では 1980 年代から研究評価における利用が図られてきたのに対し、日本ではその適用数は極めて限られていた。

➤ CHIリサーチ社(CHI Research Inc.)

米国のコンサルティング・調査会社。特許や科学文献に基づく分析やデータベースの作成、各種コンサルティングを主な業務としている。設立(1968年)当初は、科学技術文献の引用分析を主な業務としていた。特に、全米科学審議会(U.S. National Science Board)が1972年にはじめて発行した「科学指標」報告書のために、NSF(全米科学財団)とともに計量文献学的なデータに基づく国レベルの科学指標を開発したことは、その後、科学技術政策に計量文献学的分析が活用される契機となった。

1980年代には、引用分析の手法を特許情報に適用し、米国特許の引用情報を含むデータベースを開発した。特許の被引用度データや、科学と技術の連関の指標である「サイエンスリンケージ指標」は、同社が開発したものである。

➤ CRADA(Cooperative Research and Development Agreements: 共同研究開発契約)

連邦研究機関と非連邦機関(民間セクター)が共同研究開発を実施する際、技術移転関連の法律によって、共同研究で生じる知的財産権の扱い等に関する共同研究開発契約(CRADA)を結ぶことが認められている。CRADA という語は、この契約の名称であるとともに、この制度やメカニズムを指す。

CRADA は、1986 年の連邦技術移転法 (FTTA: Federal Technology Transfer Act) によって導入された。連邦技術移転法は、政府研究機関のうち政府管理型の機関 (GOGO) に対して、企業との間で共同研究開発契約 (CRADA) を結び、共同研究で生じる知的財産権に関して事前に契約することを認めた。また、1989 年の国家競争力技術移転法 (NCTTA: National Competitiveness Technology Transfer Act) により、契約者管理型の連邦研究機関 (GOCO) にも CRADA を組むことが認められるようになった。さらに、国家技術移転振興法 (1995 年) では、CRADA を結ぶ際に、企業 (非連邦機関) に対し用途限定の独占的実施権を取得する権利が保障された。なお、CRADA を通じた共同研究開発において政府機関が提供するのには研究所、生産施設、人的・施設・機器などのリソースであり、資金は提供しない。

元来、CRADA は共同研究開発行為を通じた技術移転のメカニズムとして導入されたが、最近では、民間の技術レベルが高いため、双方の資源を利用しあうコンソーシアム (リソースシェアリング) の手段といった意味合いが強くなっている。

➤ GPRA (Government Performance and Results Act: 政府業績・成果法)

クリントン政権下の 1993 年に成立した米国の法律。通常「ジブラ」と発音される。GPRA は連邦政府の施策執行機関の全てに対して組織の目的や政策目標を掲げさせ、その達成度合いを継続的に計測し、公表することを義務づけている。連邦政府各機関のミッション、戦略、具体的な業績目標を体系的に結びつけるとともに、業績目標に対する実績結果を予算配分に反映させることが狙いであり、規則遵守の行政から業績・結果重視の行政管理へ転換の試みであると評されている。

より具体的には、連邦政府機関は、少なくとも 5 年間にカバーする「中期戦略計画 (Strategic Plan)」、各会計年度における具体的な達成目標を示す「年次業績計画 (Annual Performance Plan)」、年次業績計画に盛り込まれた各会計年度における目標の達成度合いを取りまとめた「年次業績報告書 (Annual Performance Report)」をそれぞれ用意し、議会に提出しなくてはならない。

GPRA は連邦政府の様々な領域の行政管理を対象としているが、研究開発プログラム／プロジェクトという業績・結果重視の管理が困難な対象にそれを適用することを要求したという点で、科学技術関連行政は特に大きな影響を受けた領域の一つと言える。

➤ ヒンドサイト(hindsight)

hindsight という語は、「あと知恵」、「事後の判断」といった意味の一般的な英単語であるが、本報告書では「事後の洞察」といった意味で用いている。なお、hindsight という語は科学技術政策に関する言説でよく用いられるフォーサイト(foresight)という語の対語でもある。科学技術政策におけるフォーサイト(foresight)は、技術予測(technology foresight)やイギリスの Foresight Programme などに見られるように、適切な政策を立案するための将来を予見する試みを指すことが多いが、この場合、単なる予想でなく、将来を洞察する、というニュアンスがある。そのため、本報告書における hindsight という語も、単に過去の状況を見るというだけでなく、それを通じて教訓を得る、といったニュアンスを持っている。本書で紹介されている TRACE もヒンドサイトの一種である。

➤ 内部評価(internal evaluation)／外部評価(external evaluation)

本報告書においては、科学界の内部での評価を内部評価、科学界の外からの評価を外部評価と呼んでいる。この区分に依拠すると、米国の研究評価の主流であったピアレビューは科学界内部の評価であり、一方、近年、盛んになってきている定量的評価は科学界の外にも開かれた評価である、と整理できる。最近の米国の研究評価は、外部評価を取り入れる方向へ向かっている、ということができる。

なお、外部評価、内部評価という語は、これ以外にも様々な意味で用いられている。とくに我が国では、組織(あるいは「主体」)の内部、外部という区分に基づいて、内部評価、外部評価を区別する考え方が主流である。例えば、『国の研究評価に関する大綱的指針』(平成13年11月28日)では、「評価の公正さを高めるために、評価実施主体にも被評価主体にも属さない者を評価者とする外部評価を積極的に活用する。」としており、評価者の所属によって内部評価と外部評価を区別している。また、「評価規範が意思決定主体により定められるものが『内部評価』、外部から定められるものが『外部評価』である。」(市川惇信)という考え方もある。

➤ 全米アカデミー(The National Academies)

全米アカデミー(「全米アカデミー連合」と訳す場合もある)は、全米科学アカデミー(NAS: National Academy of Science)、全米工学アカデミー(NAE: National Academy of Engineer-

ing)、医学機構(IOM: Institute of Health)という三つのアカデミーから構成された組織である。NAS は議会憲章により創設され、科学技術の諸問題について、科学技術の促進と公共福祉への利用という立場から連邦政府に助言することが使命であるとされている。また NAE と IOM は NAS の綱領のもとで設立されたが、それぞれ独自の運営権を与えられている。

NRC 自体は提言活動を行う組織であり、三つのアカデミーの会員を中心に構成され、3000 名強のアカデミー会員と約 900 人のリサーチ・スタッフ(ほとんどがフルタイム)を擁している。NRC は行政府各機関とのあいだで、科学技術政策に関するレビュー契約を結び政策形成に寄与すると共に、議会からの要請により、あるいは NRC 独自の活動としても政策提言、勧告、評価を行っている。

➤ ONR (Office of Naval Research)

ONR(海軍研究局ないし海軍研究事務所と訳される)は、米国海軍省に属する科学技術研究計画の調整・実施機関であり、また基礎研究に資金提供する「助成エージェンシー」のひとつでもある。特に、基礎研究のサポートを行う連邦政府機関としては最も古い機関として知られている。下部機関には、様々な分野の基礎研究を行う海軍研究所(NRL)がある。

ONR の技術評価局長を長年務めた Ronald Kostoff は、本報告書でもとりあげているように、研究評価に関する多くの論文・著書を著していることで有名であり、それらによって ONR の研究評価への取り組みの一端を知ることができる。

➤ ピアレビュー(peer review)

同じ研究領域に属する専門家による研究評価。研究の評価はその分野の専門家(ピア=同僚)でないと判断できない、という考え方がピアレビューの妥当性の根拠となっている。ここでいう「評価」は、研究終了後に行われる「事後評価」だけでなく、研究提案の審査や選定(事前評価)、進行中の研究プロジェクトのモニタリング(中間評価)も含んでいる。例えば、我が国における科学研究費補助金(科研費)の審査もピアレビューである。

➤ プロジェクト(project)

プロジェクトは計画、事業などを意味する語である。一方、「研究プロジェクト」(ないし研究開発プロジェクト)は、経常的に行われる研究ではなく、ある目的のもとに期間を定めて実施される個別の研究を指すのが普通である。また、経常的な研究組織の枠組みを越えて特別

に編成される比較的規模の大きい研究を指す場合が多い。

研究プロジェクトは、研究上の目的や基本的な研究アプローチが事前に明確になっている場合が多いため、評価自体の困難さは他と同様であるものの、評価の基準(クライテリア)の設定は比較的容易であると考えられる。経常的研究に関しては最近まで明示的な評価が盛んでなかった我が国の公的研究機関・制度においても、研究プロジェクトの評価は実施されるが多かった。

➤ プログラム(program)

研究(開発)プログラムは、一般的には、ある政策目的を実現するための具体的な研究開発施策である。普通は、ある研究プログラムの下で個別事業として個々の研究プロジェクトが実施される。個々の研究プロジェクトが研究開発の内容面での目標を持つことが多いのに対し、研究開発プログラムは、研究内容に関する目標を設定するよりも、得られる効果、制度や運営方法の新しい試み、などの面からより長期的な目標を設定することが多い。我が国の場合、例えば「競争的研究資金制度」に分類される個別の制度は、研究プログラムである。

なお、科学技術政策に限らず、一般の政策科学や政策評価についての議論でも、「政策」、「プログラム」、「プロジェクト」という階層を考えることが多い。したがって「プログラム」は、政策と具体的な個別プロジェクトを結びつけるものと位置づけられる。これらの評価について、プロジェクト評価は公共事業評価をはじめ多くの試みがなされてきたのに対し、プログラム評価はまだ一般には十分認知されていないと言われている。

日本でプロジェクトと呼ばれているものの中には、本来プログラムと呼ぶべきものも少なくないことに留意する必要がある。