

Title	科学技術活動の計量と研究評価 : (その1) 計量書誌学的指標と研究評価
Author(s)	富澤, 宏之
Citation	年次学術大会講演要旨集, 13: 33-38
Issue Date	1998-10-24
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/5647
Rights	本著作物は研究・技術計画学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Science Policy and Research Management.
Description	一般論文

○富澤宏之 (科技庁・科学技術政策研)

科学技術活動の計量は、STS（科学技術論研究）の主要領域のひとつとして学問的見地から多様な研究が行われているだけでなく、実的な応用を目的として様々な試みが行われてきた。特に、計量書誌学的手法による定量的指標は、それがはらむ問題点ゆえに多くの論争的となりつつも、欧州のいくつかの国では、研究評価のために必要不可欠なものとなりつつある。本研究では、最近の欧米諸国における計量書誌学的手法の研究評価への適用事例およびそれらをめぐる様々な議論について、各種論文やレポート等に基づいて概観する。また、研究評価論および科学技術指標論の観点から計量書誌学的指標の研究評価への適用に関する諸問題を整理し、今後の課題を展望する。

1. 研究評価をめぐる欧米の最近の動向

計量書誌学的手法の研究評価への応用について述べるに先立ち、その背景として、欧米における研究評価全般の動向について概観したい。研究評価は近代的な「科学」に不可欠な要素であり早くから行われてきたものの、事後的研究評価が科学技術の政策手法として本格的に取り入れられるようになったのは、欧米においても比較的、最近のことである。ここでいう政策手法としての研究評価とは、研究資源の配分、それも個別研究テーマごとの研究費でなく機関レベルでの研究助成金などの配分、あるいは研究機関の廃改編のような政策上の問題のために用いられるものを指している。このような個別レベルよりも上位で研究成果を中心に行われる研究評価の導入は、1990年代における科学技術政策の特徴とでもいえるものである。このような政策手法としての研究評価に対する考え方には、欧米においても異なったいくつかの流れが見られるが、以下では、はじめに代表的なものとしてイギリスとオランダの例を示し、続いてそれらと比較しつつ他の国の状況を概観したい。

イギリスでは、1970年代の経済の低迷とともに大学の研究資金が不足し、79年頃から自国経済の競争力強化に結びつく研究の必要性が叫ばれるようになり、特定の研究分野または特定の研究機関に研究資源を集中させる重点化策（the policy of Selectivity）が本格的に取り入れられるようになった。86年に、大学の一般研究費の助成機関であるUGC（大学補助金協議会）は、大学の一般研究費の配分を、学生数に基準を置く方式から subject（日本の学科ないし大学院の専攻に相当）単位の研究評価に基づく配分方式に改めた。その後も政府主導で、研究評価システムの整備が続いており、たとえばOST（科学技術院）は、各省庁および各リサーチカウンシルに対し業績指標を作成することを求めている。これは画一的な指標を作成するというよりは、評価を各機関に行動規範（ethos）として位置づけることを意図したもの、とのことである^[1]。

イギリスの研究評価のスタイルは、しばしば「英国モデル」と呼ばれる^{*}。その特徴は、以下のよう
に要約できる。第一に、評価結果に伴う行動が事前に決められた「行動に組み込まれた評価」であ
る点を指摘できる。大学の一般研究費の配分方法に見られるように、評価結果が機関レベルでの研究
費の決定につながるのである。その背景として、イギリスでは公的歳出抑制が強く進められたきたた
め、“Value for Money”の思想が強く、競争的な環境が形成されていることを指摘できよう。第二の
特徴は、評価を行う組織の構成が「分散内在型」となっていることである。すなわち、各省庁、各リ
サーチカウンシルの内部に評価を担当する部署が設けられている。フランスなどの「集中外在型」、
すなわち政府から独立した少数個の機関が評価を行う体制と対照的である。第三の特徴は、イギリス
の研究評価は、著名な研究者や有識者ではなく、「評価の専門家」“professional evaluator”によるシ
ステマティックな評価が中心的役割を果たしていることである。なお、このような「専門家」は、評
価の専門家であっても、研究内容に関しては専門的研究者ではないのであり、研究評価に際して定量的
指標が用いられる重要な要因となっている。

以上のような特徴は、重点化政策に伴う必然ともいうべきものであり、最近では technology
foresight（日本で行われている技術予測手法を拡張したような手法）の結果を科学技術政策の重点化
に反映させるような、新しい政策手法を生むに至っている。

一方、イギリスと対照的なモデルは、オランダの研究評価に見ることができる。オランダの研究評
価の特徴を表すキーワードは“self-improvement”であるといわれている^[2]。すなわち、評価対象を直
接的に管理することではなく、評価対象が自らを向上させる手助けをすることが重視される。4 年な
いし 5 年ごとに行われる大学の評価の場合、各大学による自己評価を経て評価委員会による評価が行
われるが、そこでの評価結果は公表前に大学に示され、（積極的な意味での）交渉を行った上で公表
される。このことにより self-improvement が行われる。評価結果にもとづく予算のカットなどは行
われないが、公表されることによって学生が大学を選択する際の基準となるなど、間接的な影響は大
きいようである^[2]。このような評価システムは、最近になって整備され、1993 年以降、VSNU（大学
連合）が主導して進めている。さて、このような研究評価では結果よりもプロセスが重要となり、イ
ギリス型の評価に比べると定量的指標の重要性は少ないのではないかと考えられる。しかし、B.J.R.
Van der Meulen によると、オランダにおいて計量書誌学的指標は、各種の評価委員会で日常的に用
いられているとのことである^{[3],[4]}。ただし 1980 年代は、そのような状況にあったわけではなく、定量的
手法が受け入れられるかどうかは、指標の示す内容が評価委員会に受け入れられるかどうかによる
面が大きい、と述べている。

以上の二つの国以外でも、特に欧州では、1980 年代に多くの国で研究評価システムの整備が進め
られた。フランスの研究評価は、代表民主制の考え方を基礎としたギャランターモデルと呼ばれる方

^{*} あるいは、イギリスの科学技術政策全般についても、特徴のあるものとして、このように呼ばれることがある。

法（政策形成の方法でもある）がとられている。これは、社会を構成する各界の代表者（ギャランター）からなるパネルにより評価が行われるもので、各界の代表者は専門家でないため、代表者が推薦する専門家が支援する。

EU では、研究プロジェクトの集まりであるプログラムの評価が行われている。EU の研究開発政策は、フレームワーク計画と呼ばれる 5 年間の包括的プログラムを中核としており、このプログラム評価は、政策評価でもある。評価は独立した外部専門家グループによって行われ、外部コンサルタント調査や計量書誌学的手法の導入などの改善が続けられている^[5]。

米国では、クリントン政権の発足した 93 年に政府業績成果法（GPRA; Government Performance and Results Act）が制定され、各省庁が実施する政策プログラムの目標設定と成果測定方法の確立が義務づけられた。それ以降、科学技術分野も含め連邦政府の政策評価が盛んに行われている。具体的な評価方法については、レポート等によりいくつかの基準が示されている。米国では、個別研究テーマの評価を中心に伝統的にピアレビューが盛んであるが、ここでの評価手法はメリットレビューが中心である^[6]。

2. 計量書誌学的手法の研究評価への応用をめぐる動向

1980 年代の欧州では、研究評価が科学技術政策のなかに組み込まれる動きがある一方で、もう一方で定量的な指標に関する様々な試みと活発な議論が行われた。なかでもイギリスの SPRU（サセックス大学科学政策研究ユニット）の研究者であった B.Martin と J.Irvin の定量的指標を用いた研究評価の試みは大きな議論を巻き起こした。その評価手法は部分指標群収斂法（converging partial indicators method）と命名されており、科学研究の貢献度を測る絶対的尺度は存在しないことは認めるが、貢献度を反映する様々の部分的指標（partial indicators）が存在する、という考えに基づいている。部分指標として、(1)発表論文数、(2)発表論文の被引用回数、(3)鍵論文（特に頻繁に引用される論文）の数、(4)評価対象と同一の研究分野の専門家による評点（ピアレビューに相当する）、の四つを用い、どの部分指標をとっても首尾一貫性のある傾向が見られたとき、それを評価結果とする、というものである。実際の評価にあたっては、同一分野の研究を行う複数の研究機関を対象とし、ある一定期間での研究成果の生産性を比較している。Martin らが 1983 年に発表した論文では、電波天文学における欧州の四つの代表的研究施設を対象として、1969 年から 78 年の 10 年間について四つの部分指標を調べ、どの指標についても一貫した序列関係が見られたと報告している^[7]。

このように定量的指標によって明確な序列関係を示したために、Martin らの方法は、評価対象となった当事者からの批判を受け、また、その後 10 年以上に渡って議論が続くこととなった。しかし Martin らの試みは、特に欧州において研究評価に定量的指標が用いられる流れを作り出した要因の一つであることは確かである。その方法自体は、原理的にも実際のデータの妥当性という点でも様々な問題はあつたものの、機関レベルでのある特定の専門領域内での評価に限れば、ピアレビューのみに

よる評価を補完する方法として一定の有用性を持つことを認める関係者は多いようである。

定量的指標の評価への適用は、以上に述べたような比較的単純なものから次第により複雑なものに発展している^[8]。具体的には、計量文献学的指標の示す内容が直ちに評価結果につながるのではなく、例えば研究活動のマッピングなどと呼ばれる諸手法によって評価対象の研究活動を分析するために用いられ、評価はそれに基づいて行われる。このような定量的指標の使用方法は、次節で論じるように、評価論ないし指標論の立場からすると適切なものであるといえる。ただし、計量書誌学的方法自体にも、学際的な比較が困難であるなど、多くの課題が残されている。

3. 評価論と指標論からの考察

次に、主として研究評価論に基づいて、計量書誌学的手法の研究評価への応用についての諸問題を検討する。研究評価に関する理論のひとつとして、ここではシステム論に基づいた市川の評価論を援用する^[9]。研究評価に対するシステム論アプローチはいくつかあるが、以下での議論は

市川の評価論においては、暗黙的評価と明示的評価の区別、すなわち評価の明示性という概念が重要である。人間の活動の基本構造を計画・実行・評価サイクル（plan-do-see cycle）と見る立場から、評価とは人間が行う活動のすべてに普遍的に存在するものであり、また、何らかの形で意思決定と関係しており、評価主体の主観によって最終的に決定されるものとする。したがって評価を行っていないように見える活動も、暗黙的には評価を含んでいる。暗黙的評価では意思決定主体と評価システムが一緒になっており、評価システムが外（意思決定者以外）からは見えない形になっている。一方、意思決定主体と評価システムが分離され、評価システムが誰からも見ることができると明示的評価である。その場合、評価対象が開かれたものとなる、すなわち評価対象の集まりに自由に出入りすることが可能となるとともに、評価規範が具体的に書き出されているので、それと評価システムが整合的であるか、あるいは評価規範自体が常に再検討されることとなる。このように評価を明示的にすることにより、意思決定主体だけでなく、評価対象にも役立ち、さらにはそれ以外の関係者にも役立つようになる。

以上のような評価論から、研究評価に定量的な評価視点を取り入れる意義は何であるかを検討してみよう。評価に定量的指標を用いることの意義としてしばしば指摘されるのは、評価結果が評価者によって意図的あるいは政治的にゆがめられるのを防ぐことである。適切な定量的指標を用いることによってこのような問題を低減できることは確かである。ただし、定量的指標によって「評価の客観性」が獲得される、と考えるべきではない。市川の評価論が示すように、評価とは評価者の主観に基づくものであり、厳密な意味での「客観的評価」はありえない。定量指標は、それ自体が客観的に何かを示したとしても、評価結果を直接示しているわけではなく評価規範が設定されて初めて評価がなされるのである。もちろん、定量指標が客観的に示した内容がそのまま評価結果と一致する場合もありえるが、それはたまたま評価規範が指標の示す内容と一致していたに過ぎない。市川の評価論の立場か

らすると、評価に定量的指標を用いる意義は、評価システムを暗黙的なものでなく明示的なものにするため点にあるということができよう。

評価と定量的指標との関係で、もうひとつ注意を払わなければならないのは、評価における指標の機能である。市川は評価のプロセスについて、まず評価対象のもつ様々な属性を各属性ごとの軸が形成する属性空間に位置づけ、次にそれをいくつかの評価項目ごとの軸が形成する評価空間に移し、最終的に意思決定空間に移す、という形に定式化している。さらに、このプロセスのある段階までは、できるだけ客観的に行うことが有用であること、また最終的に意思決定に結びつけるためには、属性空間および評価空間における測定値の並びがもつ性質を分析することが重要であると指摘している。

次に、以上の議論を科学技術指標論の立場から補足しておこう。指標の機能は、評価以外にも、(1)対象についての認識や理解・把握、(2)理論形成および理論検証、(3)予測、(4)制御ないし操作、(5)意思決定、(6)コミュニケーション（指標を通じた認識の共有。客観性の確保など）、などがある。これらは独立ではなく、部分的に重複したり、あるいは包含関係にあるなど相互に複雑な関係にある。特に、認識・理解・把握（簡単に「理解」としておく）という機能は、他のほとんどの機能の一部を成している。

以上をふまえると、定量的指標は、評価に伴う対象の理解、および判断材料を提供するためのツールであり、より厳密に言うならば、対象を理解し判断する思考のある部分を明示的にするものである。評価に際しては、対象が設計されたシステムでない限り、それを間接的に把握し、その範囲内でできるだけ合理的に判断を下すことしかできないが、対象を把握し、しかもそれを評価に結びつけるためには、操作性をもつ定量的指標の果たす役割は大きい^[10]。市川の言う属性空間・評価空間における分析とは、対象の理解のために行うだけでなく、対象の理解を評価のフェイズへ移行させるために必要なステップであり、定量的指標はそこで重要となる仲介機能を果たすのである。

以上の議論から一つの指摘をしておきたい。それは「評価指標」という語はしばしば不適切であるということである。指標という語は「何かを指し示すもの」を意味するので、評価指標とは評価結果を示す指標という意味になる。確かに評価結果を示す指標も前述のように存在するが、以上の議論が教えるように、評価に必要となる指標はそれだけではない。言い換えるなら、評価に用いられる指標には大きく分けて二つあり、ひとつは指標自体が評価規範を含んでおり評価結果を示す指標であり、もうひとつは評価対象についての認識を深めるなど評価のプロセスで重要となる指標である。

4. まとめ

日本では研究評価が本格的に行われるようになったのがごく最近であることもあり、「日本は定量的な評価アプローチが利用されていない最後の国のひとつ」^[11]という論評は極端であるとしても、現時点では定量的手法がほとんど行われていないことは確かである。その適否については本稿の範囲を超える問題であるが、使い方次第では有用なツールとなる可能性を持っていると考えられ、一層の研

究を行う意義は大きいであろう。前節までに述べたように、定量的指標は評価を明示的にするために用いることができるのであり、決して客観的な評価尺度を提供するわけではない。また、Van der Meulen が指摘したように、定量的指標の使用は、指標の示す内容を評価者（あるいは他の関与者）が納得した場合に受け入れられる。これを評価論の立場から解釈すると、明示の評価は暗黙の評価と整合する場合に受け入れられるということである。このようなことから、評価の明示化、という観点からの検討は特に有効であろう。

参考文献

- [1] 館和夫, 「英国における研究評価 ―公的研究にみる評価 “Value for Money” と “Selectivity”―」 (科学技術政策研究所／調査資料・データ No.54), 科学技術庁科学技術政策研究所, 1998 年 5 月.
- [2] 財団法人政策科学研究所, 「科学技術の戦略的な推進に関する調査 ①海外主要国の科学技術政策形成体制の動向調査」 (平成 9 年度科学技術振興調整費／調査研究報告書), 1998 年 3 月.
- [3] B.J.R. Van der Meulen, “The Use of S&T Indicators in Science Policy: Dutch Experience and Theoretical Perspectives from Policy Analysis”, *Scientometrics*, Vol.38, No.1, January 1997.
- [4] Jan van Steen and Marcel Eijffinger, “Evaluation practice of scientific research in the Netherlands”, *Research Evaluation*, Vol.7, August 1998.
- [5] P.Cunningham, “The Evaluation of European Programmes and the Future of Scientometrics”, *Scientometrics*, Vol.38, No.1, January 1997.
- [6] R.N.Kostoff, “Federal Research Impact Assessment”, *Scientometrics*, 34, 163-206, 1995.
- [7] B.R.Martin and J.Irvine, “Assessing Basic Research: Some Partial Indicators of Scientific Progress in Radio Astronomy”, *Research Policy*, 12, 1983.
- [8] Yoshiko Okubo, “Bibliometric Indicators and Analysis of Research Systems: Methods and Examples”, *STI Working Papers 1997/1*, OECD, 1997.
- [9] 市川惇信, 「研究評価について」 (科学技術政策研究所／研究評価論講演シリーズ／講演録-46), 科学技術庁科学技術政策研究所, 1996 年 11 月.
- [10] 富澤宏之, 「科学技術指標へのシステムズ・アプローチ: 人間活動システムの測定論」, 研究・技術計画学会, 第 12 回年次学術大会講演要旨集, 259-264 (1997).
- [11] David Swinbanks, “Statistical test rates research strengths.” *Japan Times*, November 13, 1994.