

# 米国における研究評価

—NSF, NIHの例—

長田 洋

## 1. はじめに

米国では、研究評価においてピア・レビュー (Peer review) とよばれる方法が古くから採用されている。このピアとは元来“同僚”であり、第三者の専門家による評価をさしている。この方法は学会誌の投稿論文の審査に用いられたレフェリー評価に由来し、基礎研究の申請書の評価、つまり事前評価に多く使用され、効果を発揮している。米国の政府機関ではNSF (National Science Foundation, 国立科学財団), NIH (National Institutes of Health, 国立衛生研究所), 国防省などのピア・レビューシステムが著名である。このようなシステムはわが国の文部省の科学研究費の配分における評価にもみられる。本稿では世界的にも最も洗練され、各国の研究評価の参考にされているNSFとNIHのピア・レビューを紹介する。

## 2. NSFのピア・レビュー

### (1) 概要

NSFは独立した連邦機関で米国における科学・工学の促進と発展を目的として1950年に設立された。NSFの研究助成はあらゆる分野にわたり1980年度の研究補助金額は6億5300万ドル、1981年度は9億5550万ドルであり、これは米国政府の

基礎研究投資額の18%を占めている。NSFの助成先は2000を超える大学が約70%を占め、以下国立研究所、非営利機関となっている。NSFへの助成応募件数は1980年で26,000件、採用率は44%、平均補助金額は1件当たり56,800ドルであった。助成研究の平均期間は約2.2年であった。

### (2) 申請書の評価

NSFへの補助金申請は、通常、希望する研究開始日の9~12カ月前になされる。申請書は経歴、研究プロジェクトの概要、研究予算、現在の助成状況、最近の業績などを記入した資料からなっている。提出された申請書は次のいずれかの方式で評価される。

#### ① 郵送による評価

化学、物理・数学部門では申請書はその分野の提案者に利害関係のない専門家 (ピア) に郵送され、評価をうける。ピアは3~8人 (またはそれ以上) の科学者である。

#### ② パネル郵送による評価

地球科学部門、生物・行動科学部門、社会科学部門の研究は郵送による評価の他にピアを数人集めたパネルによる評価をうける。パネルは3~7日間、缶詰になり、評価の討論が最終的にある程度のコンセンサスが得られるまで行なわれる。

いずれの評価でも評価基準は、

① 申請された研究の科学的意義——特に独創性と創造性

② 主任研究員 (プロジェクトリーダー) の最近

おさだ ひろし (旭リサーチセンター)

の研究業績——研究遂行能力が推測される

③予算の妥当性と設備能力——要求金額の妥当性と研究遂行に必要な設備のチェック

などである。そしてどの申請書も「Excellent」, 「Very Good」, 「Good」, 「Fair」, 「Poor」の5段階で評点がつけられる(図1)。さらに評価者によるコメントが付与される。次にNSFのプログラムオフィサーが、評価者のコメントを重視したうえで、全体的な最終評価を行ない、申請書に対して資金供与の意思決定を下す。

なお評価内容については申請者が希望すれば、すべて公開される。この点は後に述べるNIHでも同様であり、非公開が多い日本の評価システムと異なる点である。助成がなされなかった研究者に対しては評価者のコメントが今後の研究に非常に役立っているといわれる。

### (3) プログラムオフィサー

プログラムオフィサーは評価者の選定者であり、実質的な予算決定者である。彼らはその担

分野に研究経験を有する科学者、エンジニアである。プログラムオフィサーには常任スタッフと1~2年の交替制で任期終了後に所属機関に復職する者がある。

### (4) 評価者の選定

各分野での評価者はプログラムオフィサーにより、次の基準にもとづいて選定される。

①主任研究員の能力とプロジェクトで利用される設備やデータを評価できる人

②申請書の主題およびそのプロジェクトの成功により影響をうける分野に関して識見を有する人

③そのプロジェクトにより影響をうける科学界の組織、教育面などを代表する人

このような基準を満たす評価者は分野別にファイルされている。またプログラムオフィサーは常に学会などと連絡を密にして評価者グループの充実を図っている。

### (5) 評価へのインセンティブ

このようにして選ばれた評価者は多忙な時間を

NATIONAL SCIENCE FOUNDATION		PROPOSAL EVALUATION FORM		NSF Form 18 (9-81) Supersedes All Previous Editions
PROPOSAL NO.	INSTITUTION			PLEASE RETURN BY
PRINCIPAL INVESTIGATOR		NSF PROGRAM		
TITLE				
COMMENTS (QUALITY OF THE PROPOSED RESEARCH, RECENT RESEARCH ACHIEVEMENTS OF THE PRINCIPAL INVESTIGATOR(S), ETC.) CONTINUE ON ADDITIONAL SHEET(S) AS NECESSARY.				
OVERALL RATING: <input type="checkbox"/> EXCELLENT <input type="checkbox"/> VERY GOOD <input type="checkbox"/> GOOD <input type="checkbox"/> FAIR <input type="checkbox"/> POOR				
Verbatim but anonymous copies of reviews will be sent only to the principal investigator/project director. Subject to this NSF policy and applicable laws, including the Freedom of Information Act, 5 USC 552 and formal requests from Chairpersons of Congressional committees having responsibility for NSF, reviewers' comments will be given maximum protection from disclosure.				
REVIEWER'S SIGNATURE		REVIEWER'S NAME (TYPED)		
OTHER SUGGESTED REVIEWERS (OPTIONAL)				
PROPOSAL FILE				

図1 NSFの評価表

割き、評価に協力するにもかかわらず、謝礼は一切ない。では何のために彼らは評価するのか？

彼らにとってNSFのピア評価者に選ばれることは、その分野の専門家として認められたことを意味している。実際、評価者のリストは毎年出版されており、公表されることにより、専門家としての地位が高まるのである。と同時に評価は専門家としての責任であると見なされるのである。

第2の評価へのインセンティブは新しい研究上の知識の習得である。他人の新しい研究を評価するには、その研究内容を理解しなければならず、そこにおける意見や方法は彼らの研究への刺激となり、間接的には知識となるのである。

### 3. NIHのピア・レビュー

#### (1) 概要

NIHは国立ガン研究所、国立環境衛生科学研究所など10の研究所、国立医学図書館、国際センター、6つの研究サービス部、事務局から構成されている世界でも最大かつ最先端のライフサイエンスの研究機関として知られている。

NIHの1980年度の研究補助金総額は約16億ドルであり、助成応募件数は約26,600件、採択率は62%であった。申請1件当りの平均助成金額は97,000ドルであり、平均研究期間は3.2年といずれもNSFの規模を上回っている。

NIHは研究評価を重視しており、評価システムの運用・開発など評価に対し、研究開発費の1.5~2%のコストを割いている。

NIHにおけるピア・レビューもNSFと同様に申請書の評価において最も効果的なものとなっている。

#### (2) 申請書の振り分け

研究計画の概要、予算(5年間)、主任研究員およびプログラム管理者の経歴、他の研究助成状況、研究資源(設備など)や環境などが記載された研究補助金申請書はNIHの研究助成部へ送られる。そこで科学専門官が申請書を読み、その申

請書の専門領域に関する専門家がいるスタディ・セクション(Study Section)へ申請書を回す。もし既存のスタディ・セクションが適切でない場合は、新たに特別なスタディ・セクションが設置される。スタディ・セクションでは専門的評価がなされる。NIHの科学専門官は5あるいはそれ以上のスタディ・セクションを含む1つのまとまった専門分野を担当しており、スタディ・セクションへの振り分けに重要な役割を担っている。

#### (3) スタディ・セクションでの評価

NIHのピア・レビューの評価の特徴は2段階評価(Dual Review)にある。すなわち第1段階はスタディ・セクションでの専門的評価であり、第2段階は国家諮問委員会による非専門的・総合的評価である。

スタディ・セクションは現在、表1に示すように63の専門領域に対応してそれぞれ設置されている。スタディ・セクションの構成メンバーは10~15人の有能な専門家(コンサルタント)であり、大学や国立研究所の研究者が多い。彼らは申請書の評価者となるが、それぞれの専門領域で有能かつ高い業績を有することがメンバーとしての基準であり、スタディ・セクションのメンバーに選ばれることはライフサイエンスでの専門家として認められたことを意味しており、名誉なことでもある。任期は4年であり、再任されない。メンバーリストは年2回、公表されている。また毎年メンバーの1/4ずつが入れ替わっている。

さて、各スタディ・セクションにはエグゼクティブ・セクレタリ(Executive Secretary)とよばれるスタッフがあり、彼が実質的な評価の責任者である。エグゼクティブ・セクレタリは申請書を評価者であるスタディ・セクションのメンバーに送付し、申請書に対する評価・コメントを求める。コメントが返送されると6~8週間後にエグゼクティブ・セクレタリによりスタディ・セクション会議が招集される。

スタディ・セクション会議は年3回開かれ、期

間は1回が2~3日間である。

1回の会議で約50~100の申請書が個々に評価され、かなりハードな会議といえる。

会議ではまず2人以上の選ばれたメンバーが申請書を検討し、その結果を文書にして他のメンバーに公表・説明する。次に全員による討議がなされる。そして次のような評価基準にしたがって、申請書に対する承認、否決、保留の投票がなされ、多数決により承認された申請書について、各メンバーは主として科学的見地から評点をつける。(1.0(優)~5.0(不可))

(評価基準)

- ①申請された研究の学問的価値と独創性
- ②主任研究員の経験、能力
- ③計画の妥当性
- ④必要な施設や研究資源の有無
- ⑤申請された予算や研究期間の妥当性

評点は平均され100倍される。平均評点はコンピュータに入力され、他の2回の会議の結果と、評点の分布(平均値、標準偏差)が同一になるよう変換され、優先順位づけがなされる。この優先順位が最終決定の際に効力を発揮する。

スタディ・セクション会議の後、エグゼクティブ・セクレタリは各申請書についての

表1 スタディ・セクションの種類(NIH)

CODE	STUDY SECTION
ALY	ALLERGY AND IMMUNOLOGY
BM	BACTERIOLOGY AND MYCOLOGY
BEM	BEHAVIORAL MEDICINE
BNS	BEHAVIORAL AND NEUROSCIENCES FELLOWSHIPS
BMT	METALLOBIOCHEMISTRY
BCE	BIOCHEMICAL ENDOCRINOLOGY
BIO	BIOCHEMISTRY
BI	BIOMEDICAL SCIENCES FELLOWSHIPS
BNP	BIO-ORGANIC AND NATURAL PRODUCTS CHEMISTRY
BBCA	MOLECULAR AND CELLULAR BIOPHYSICS
BBCB	BIOPHYSICAL CHEMISTRY
BPO	BIO-PSYCHOLOGY
CVA	CARDIOVASCULAR AND PULMONARY
CVB	CARDIOVASCULAR AND RENAL
CBY	CELL BIOLOGY
CPA	CHEMICAL PATHOLOGY
CLN	CLINICAL SCIENCES FELLOWSHIPS
CMS	COMMUNICATIVE SCIENCES
RNM	DIAGNOSTIC RADIOLOGY
END	ENDOCRINOLOGY
EDC	EPIDEMIOLOGY AND DISEASE CONTROL
ECS	EXPERIMENTAL CARDIOVASCULAR SCIENCES
EI	EXPERIMENTAL IMMUNOLOGY
ET	EXPERIMENTAL THERAPEUTICS
EVR	EXPERIMENTAL VIROLOGY
GMA	GENERAL MEDICINE A
GMB	GENERAL MEDICINE B
GEN	GENETICS
HEM	HEMATOLOGY
HUDEV	HUMAN DEVELOPMENT AND AGING
HED	HUMAN EMBRYOLOGY AND DEVELOPMENT
IMB	IMMUNOBIOLOGY
IMS	IMMUNOLOGICAL SCIENCES
MGN	MAMMALIAN GENETICS
MCHA	MEDICINAL CHEMISTRY
MET	METABOLISM
MBC	MICROBIAL PHYSIOLOGY
MBY	MOLECULAR BIOLOGY
CTY	MOLECULAR CYTOLOGY
NLS	NEUROLOGICAL SCIENCES
NEUA	NEUROLOGY A
NEUB	NEUROLOGY B
NTN	NUTRITION
OBM	ORAL BIOLOGY AND MEDICINE
ORTH	ORTHOPEDICS AND MUSCULOSKELETAL
PBC	PATHOBIOLOGICAL CHEMISTRY
PTHA	PATHOLOGY A
PTHB	PATHOLOGY B
PHRA	PHARMACOLOGY
PB	PHYSICAL BIOCHEMISTRY
PC	PHYSIOLOGICAL CHEMISTRY
PHY	PHYSIOLOGY
RAD	RADIATION
REB	REPRODUCTIVE BIOLOGY
SSP	SOCIAL SCIENCES AND POPULATION
SSS	SPECIAL PROGRAMS
SAT	SURGERY, ANESTHESIOLOGY, AND TRAUMA
SB	SURGERY AND BIOENGINEERING
TOX	TOXICOLOGY
TMP	TROPICAL MEDICINE AND PARASITOLOGY
VR	VIROLOGY
VISA	VISUAL SCIENCES A
VISB	VISUAL SCIENCES B

検討・討議内容を文書にし、次の国家諮問委員会に送る。

#### (4) 国家諮問委員会での評価

国家諮問委員会はスタディ・セクションでの専門的評価にもとづき、より高い次元からの評価を行なうために設立された組織であり、そのメンバーは12名である。半数はNIHの研究プログラムに直接関係のある科学・教育分野の専門家であり、残りは国民衛生の分野で活躍する一般人である。任期は4年である。委員会はスタディ・セクション会議終了後6～8週間以内に開かれ(年3回)、次のような評価基準にしたがい、申請書の評価がなされる。

(評価基準)

- ① NIHのミッションとの整合性
- ② NIHの重点分野との関連性
- ③ 社会ニーズとの合致度およびそれにもとづく資金配分の妥当性

そして委員会終了から数週間後に申請者に助成に関する結果が通知されるのである。

#### (5) IMPAC

NIHの評価システムを支える基盤として見逃がすことができないのは研究情報管理システム、IMPACである。IMPACではNIHが助成する(した)研究プロジェクトの助成前の評価に関する情報、助成後の研究情報(研究者名、工数、コスト、進捗状況)、研究成果など、あらゆる情報がデータベース化されている。現在、約24,000のプロジェクトに関する情報が収録され、120人のスタッフがこのシステムの運用にたずさわっている。このようなシステムは事前評価のみならず中間評価、事後評価に必要な情報を提供してくれ、また評価結果のフォロー、つまり助成した研究の投資効果分析や評価法の見直しにも活用されている。

## 4. ピア・レビューの見直し

### (1) NSFでの見直し

以上紹介したようにピア・レビューは第三者評価とはいえ、究極には主観評価であり、評価者への依存度は大きい。NSFではピア・レビューの客観性、不偏性を確認し、よりよい評価システムに改善するためにピア・レビューの見直しを行なった。この見直しのための調査は全米科学アカデミー(National Academy of Science)によって、過去2回にわたって行なわれた。

#### (i) フェーズIの結果

1975年にCole教授らによって調査がなされ、1978年に調査結果が発表された。

調査はNSFのスタッフと75回にわたるインタビュー、10部門から抽出された1200の申請書の分析、さらにそのうち250の申請書に対する評価者の批評の分析からなっている。その結果、

① 評価者が与えた評価点が高いほど助成金をもらう確率が高い。

② 研究費をもらった人とその人の過去の業績とはあまり大きな相関はない。つまり過去にすばらしい業績を残した有名人でも助成金をもらえるとは限らない。

③ 有名大学の研究者と小さな大学の研究者との申請書の評価には有意差がない。

④ 研究者の経歴年数は評価に影響しない。

⑤ 研究者の所属する大学の評判や場所、過去5年間にわたるNSFからの助成状況などと評価結果とは相関なし。

などがわかり、NSFのピア・レビューは組織的な偏見がないと結論づけられたのである。

#### (ii) フェーズIIの結果

ついで第2回の調査は第1回目と同じくColeらにより1977年に開始され、1981年に結果が公表された。この調査ではNSFの評価者によって評価がどのような影響をうけるかを確認するためのものであった。そのためにまず化学、経済学、固体物理学の分野から評価がすでに終わっている申請書がそれぞれ50ずつ抽出された(半数は助成金がもたらえたものである)。次に各分野から10～18

人の専門家が選ばれ、彼らのうち2人が各申請書ごとに6人ないしそれ以上の評価者を選んだ。その結果、各申請書に対し平均約12人の評価者が選出された。

次に上記の評価者がNSFにも同一の評価基準で申請書を評価したところ、助成されるべきものと、実際に助成されたものとの一致度は70%であった。つまり同じような評価者が同じ基準にしたがって評価しても約30%は同じ結果が得られないということである。

NSFでは、この“衝撃的”な調査結果をもとに、評価者の数をふやし、上記のような再現性を高めることを検討している。

## (2) NIHでの見直し

NIHではピア・レビューの見直しのための調査小委員会を1975年4月に設置し、調査を行なった。そこでは1400人に手紙でピア・レビューに関する意見を求め、93人にヒアリングを行なった。さらに1200人のNIHのピア評価者(1975~76年)にアンケートを送付した。

その結果、約12000のコメントが寄せられたが、ピア・レビューシステムの改革、変更を訴えるものが25%であり、全面的改訂を希望したのは50人であった。約83%がこのシステムを優れていると評価した。またNIHの評価者だった研究者のうち96%が現行のシステムに賛成であった。

このような結果はほぼNIHのピア・レビューが妥当なものであることを示しているが、NIHはさらに改良をめざして不当な評価結果に対し、応募者が異議申し立てができるような制度の確立を検討し始めた。

## 5. ま と め

以上、米国の公的基礎研究に対する研究評価の代表例を紹介したが、それらは欠陥はあるが現行では最良の評価システムといえる。その評価手法としては簡単な評点法が中心である。それよりもむしろ両機関が共通して重視しているのは、客観

的・公平な評価であり、そのために、

①評価者の選定法

②評価システムの規模、形態（評価者数、評価形態—パネル、郵送など）

③評価結果の公開

④評価に必要な情報システムの整備

⑤評価システムの見直し、改良

などに腐心している。もちろん評価手法についても開発がなされ、特に研究成果の定量的把握のために論文の引用頻度(Citation Index)などを指標とする計量文献学的手法(Bibliometrics)も熱心に研究されている。

しかし、上記①~⑤のような諸項目は研究評価システムの定着、運用に欠かすことができない、いわば評価の基盤的条件であり、過去、いたずらに高度な評価手法の追求に執心し、肝心な評価システムの定着についてあまり関心が払われていなかったわが国において、大いに参考にすべき点であると思われる。

## 参 考 文 献

- [1] 旭リサーチセンター：研究評価のあり方に関する調査研究(56年度), 1982
- [2] Sanders, H. J. : Peer Review—How well is it Working? *Chemical & Engineering News* Vol. 60, No. 11 (1982) 32-43
- [3] NSF : *Grants for scientific and engineering research*. NSF, 1981
- [4] NIH : NIH Peer review of research grant applications, NIH, 1981
- [5] Cole, S. et al. : Peer Review in the National Science Foundation, Phase one. National Academy of Sciences, 1978