

3 研究開発プログラムの評価

三菱総合研究所

要 旨

国による研究開発は、政策から個別の研究開発課題まで階層的にマネジメントされており、その中でも、政策と研究開発課題・プロジェクトの中間階層に相当するプログラムレベルのマネジメントが重要視されてきている。こうしたマネジメントに対応してプログラムレベルでの研究開発評価が必要である。ただし、プログラムレベルでのマネジメント、プログラムレベルでの研究開発評価も我が国では試行錯誤が行われている段階である。

プログラムレベルのマネジメント・評価を行うことによって、研究開発の位置づけや効果の明確化、政策・戦略と研究開発課題との連動の強化、効果的・効率的な研究開発課題のマネジメント、研究開発評価の簡素化が期待できる。

各国の先行事例では、政策目標に対応したプログラムを設定して評価し、プログラムの中のプロジェクトの資源配分はプログラムに委ねるマネジメントが見られる。

I プログラム評価の背景

国費を用いて実施される研究開発は、それらをマネジメントする主体の面からみても、また、評価の対象となる研究開発の面からみても、上位である政策から、個別の研究開発課題・プロジェクトまでの階層構造となっている。評価は、このような階層構造の下で各々の階層レベルにおいて重層的に実施される⁽¹⁾。プログラムは、政策と研究開発課題・プロジェクトの中間階層に相当するものであり、それを評価するのがプログラム評価である。

「国の研究開発評価に関する大綱的指針」では、評価の対象として、研究開発課題に加えて、複数の個別研究開発課題等を政策上の特定の目的や目標ごとにひとつのまとまりとした施策、競争的資金制度、さらに、分野ごとの研究推進方針や戦略、計画などの体系を挙げている⁽²⁾。

また科学技術・学術審議会の「研究開発評価システム改革の方向性について」⁽³⁾では「政策」－「施策」－「プログラム・制度」－「研究開発課題・プロジェクト」という階層を提示し、「プログラム・制度」は次のように説明されている。なお、「プログラム・制度」は、政策評価における「政策（狭義）」－「施策」－「事務作業」の体系の「事務作業」に対応する⁽⁴⁾。

- 施策を具現化するための個々の行政手段としての事務及び事業であり、行政活動の基礎的な単位となるもの。
- 国や資金配分機関において実施される提案型研究開発公募事業などの研究開発事業、競争的資金による研究開発制度が相当する（大規模なプロジェクトであり、資金配分や運営の意思決定が単独のプロセスで行われるものはここに含まれる。）。
- 特に、科学技術政策においては、プログラムや制度は、社会、経済、国際政治、環境、

(1) 「国の研究開発評価に関する大綱的指針（別冊）」（平成20年10月31日内閣総理大臣決定）、5頁。

(2) 同上、15頁。

(3) 「研究開発評価システム改革の方向性について（審議のまとめ）」

<http://www.mext.go.jp/a_menu/kagaku/hyouka/1289630.htm> [last accessed 2012/2/7]

(4) 「政策評価の実施に関するガイドライン」（平成17年12月16日政策評価各府省連絡会議了承、平成22年5月28日一部改正）、2頁。

文化などの様々な政策的課題（学術政策の課題も含む）と、最先端の研究動向とを結びつけるインターフェイスの役割を担う点に重要性がある。

- その評価は、プログラム・制度の設定、研究開発の質の向上や運営改善、計画の見直し等を図るとともに、評価の重複を避け、評価を効率的に実施するためにも極めて重要である。

研究開発評価の階層と種類の分類について、一例を以下に示す。

図1 研究開発評価の階層と種類

階層	政策				
	施策				
	プログラム・制度				
	研究開発課題、プロジェクト				
		事前評価 (アセスメント)	中間評価 (モニタリング)	事後評価 (エヴァリュエーション)	追跡評価
		実施時期			

(出典) 科学技術・学術審議会研究計画・評価分科会研究評価部会「研究開発評価システム改革の方向性について（審議のまとめ）」(平成21年8月4日)を基に筆者作成

欧米では、政策における評価の基本単位がプログラムとなっており、個々のプロジェクトではなく、プロジェクトを束ねたプログラムを対象とした評価が実施されている。

我が国では、政策評価は府省において政策評価法に基づき実施され、研究開発課題・プロジェクトの評価は府省やファンディング・エージェンシー等で実施されているのに対し、プログラムの評価は、府省や独立行政法人などの研究開発機関の内部で実施されているものの、まだ十分な経験が積みまれておらず、その実施方法についてはなお検討の必要がある。

プログラムは、共通目的の下で研究開発課題群をまとめたものであり、個別の研究開発課題よりも実施期間は長期となる。こうしたプログラム単位でマネジメントし、評価をしていくことは以下のようなねらいが考えられる。

表1 プログラムレベルでのマネジメントと評価を行うねらい

ねらい	現状の課題	プログラム単位のメリット
研究開発の位置づけや効果の明確化	研究開発課題単位ではそれぞれがどのような関係にあり、どのような成果が期待されるのかが分かりにくく、重複や漏れが生じやすい。	期待するアウトプット・アウトカムや他の政策領域との棲み分けが明確になる。アウトカムは個々の研究開発課題の成果ではなく、それらの組み合わせによって一定の期間を経て得られるものである。
政策・戦略と研究開発課題との連動の強化	政策の次は、個別の研究開発課題となっていると、それぞれの関係がわかりにくく、連動が難しい。 政策に本来関係が薄い研究開発課題が実施されたり、研究開発課題の成果が政策・戦略にうまく活かされなかったりする。	上位の階層である政策・戦略（国家戦略・政策や事業戦略など）と、下位の階層である個々の研究開発課題との連動を強化し、研究開発の位置付けをより明確化し、実施効果を発揮しやすくする。 研究者は研究の政策上の位置付けを理解しやすく、研究成果の発現、成果の円滑な受け渡しを測るためのインセンティブが働く。プロジェクト単独では発揮しづらい政策効果への関与が明瞭になる。
効果的・効率的な研究開発課題のマネジメント	ある研究開発課題が成功しても他の研究開発課題が失敗した結果、目的が達成できない。 研究の進捗に応じて各課題への資源配分を見直せない。 ある研究開発課題の成果が他の研究開発課題で活用されない。	各研究開発課題の理解が容易となり、全体を俯瞰した進捗管理ができる。プログラムに包含される研究課題間での成果の展開、研究員相互の知識創造によりシナジーを促す。また、課題の重複を軽減し、効果的な研究開発を推進する。 プロジェクト間の相互関係をみることで、効果的なプロジェクト配置が可能になる。（不足研究課題の追加、研究の加速、縮小等の判断をしやすくなる）
研究開発評価の簡素化	細かな個別研究課題まで評価する負担が大きい。 個々の研究開発課題の成果の良否を問うてもが必ずしも政策の実現につながらない。	研究開発評価の簡素化に寄与する。政策として実効性などを見るのが重要であり、個々の研究開発課題の成果評価を厳しく行っても意味がない場合が多い。

(出典) 文部科学省「研究プログラムの構成の考え方とその評価体系、評価手法及び評価者の視点を通じた研究コミュニティの活性化に関する調査・分析報告書（研究プログラムの構成の考え方とその評価体系、評価手法に関する調査・分析編）」2010、180頁を基に筆者作成。

なお、「プログラム」に相当する階層の呼称についてはさまざまなものがあり、研究開発の中でどの階層をプログラムと呼ぶかは明確ではない場合もある。また、プログラムについては、研究開発の基礎・応用といった性格の違い、トップダウン的に設計される場合とボトムアップ的に設計される場合の違いがあることに注意が必要である。

プログラムの運営に関しては、責任者が置かれる。我が国の競争的資金の場合、プログラムオフィサー（Program Officer: PO）（各制度の個々のプログラムや研究課題の選定、評価、フォローアップ等の実務を行う研究経歴のある責任者）や、その上位職であるプログラムディレクター（Program Director: PD）（競争的研究資金制度と運用について統括する研究経歴のある高い地位の責任者）が置かれる。個々のプログラムではプログラムマネージャー等、異なった呼称も用いられることがある。

II プログラム評価の事例

1 日本

(1) 科学技術振興機構における戦略的創造研究推進事業（CREST）

戦略的創造研究推進事業は課題解決型基礎研究を行うものであり、CREST、さきがけ、ERATOといったプログラムが含まれているが、CRESTはその中で最大規模のものである。

CRESTでは、国の科学技術政策や社会的・経済的ニーズなどの政策目標をもとに社会的インパクトの大きい目標（戦略目標）を国（文部科学省）が設定する⁽⁵⁾。戦略目標は、将来の研究成果の国民への還元の色を示すものとなる。次に科学技術振興機構（Japan Science and Technology Agency: JST）では、この戦略目標にもとづき、推進すべき研究領域を定め、研究領域ごとに研究総括を選定する。

このようにCRESTは大きな研究開発プログラムとして捉えることができ、さらに、それぞれの研究領域も目的を持ったプロジェクト群であり、研究総括がプログラムオフィサーとして権限と責任を持って推進するということから1つの研究プログラムとみなすことができる。

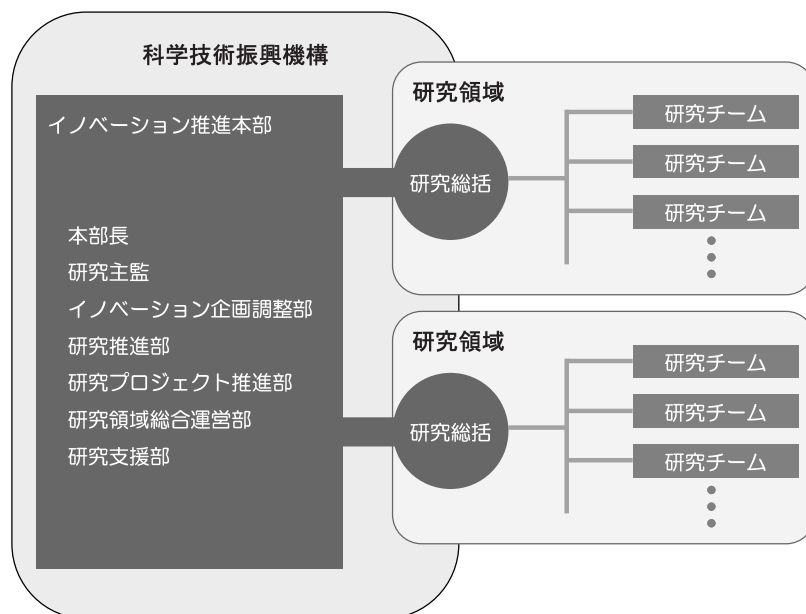
研究領域毎に研究提案の公募が行われるが、研究総括は、採択課題の決定、研究計画の調整、研究代表者との意見交換、研究への助言、課題の評価までも行っている。

概ね1研究領域は10～20程度の研究チームから成っている。各研究課題の研究期間は5年以内と定められている。

評価は時期について事前評価から追跡評価があり、階層については事業レベル、領域レベル、課題レベルの3層構造になっている⁽⁶⁾。

研究領域評価については、個々の領域の事前評価、中間評価、事後評価、追跡評価がある。また、CRESTで実施されている研究領域全体の領域総合評価も行われている。評価は外部有識者からなる評価委員会が行い、領域評価においては戦略目標達成の視点が最も重要となっている。中間・事後評価では、研究成果および戦略目標の達成状況が評価され、追跡評価では副次的効果を含めて研究成果の発展状況や活用状況等が評価される。

図2 CRESTの体制



（出典）科学技術振興機構ウェブサイト<<http://www.jst.go.jp/kisoken/crest/about.html>> [last accessed: 2012/2/7] を基に筆者作成

(5) 科学技術振興機構ウェブサイト「CRESTとは」<<http://www.jst.go.jp/kisoken/crest/about.html>> [last accessed: 2012/2/7]
 (6) 科学技術振興機構「基礎研究に係る課題評価の方法等に関する達」<<http://www.jst.go.jp/kisoken/basic.html>> [last accessed: 2012/2/7]

研究課題評価については、研究総括が領域アドバイザーの協力を得て責任を持って評価しており、中間評価結果は、研究チーム編成見直しや資源配分に反映される。

戦略目標の達成状況は、研究課題レベルではなく、研究領域レベルでの評価に重点がおかれていることがわかる。

(2) 経済産業省の施策（プログラム）

経済産業省では、IT、ナノテク・部材、ロボット・新機械、健康安心、エネルギー、環境安心、航空機・宇宙産業といった7つのプロジェクト群を作り、それぞれをイノベーションプログラム（IPG）としている。ここでのプログラムは、「政策目標を達成するために必要な、研究開発課題の設定、資源配分、成果の政策上の活用の方策、政策実現のための環境整備や他の政策との連携等まで含めたパッケージ」とされている⁽⁷⁾。これは17の研究プログラムが再編されたものである。

評価についても平成20年度末よりプログラム評価を実施しており、標準的評価項目が評価指針にも盛り込まれている⁽⁸⁾。プログラム評価の目的は平成20年度経済産業省技術評価調査「プログラム評価の手法等に関する調査」報告書によると次のようになっている。

1. 政策とプロジェクト間の距離感を埋め合わせるプログラムを設定し評価を行うことでより、プロジェクトの評価が政策と密接に関連付けられる。
2. プロジェクト間の相関関係を見ることができ、目標を達成する上における適切なプロジェクト配置が可能になる。
3. プログラム内の無駄をなくし、効率的な資金運用を図ることができる。
4. 効率的な評価を実施することができる。
5. プログラムのマネジメントを評価する。

プログラム評価の実施によって、従来プロジェクトごとにばらばらに中間評価や事後評価を実施し、毎年評価検討会および産業構造審議会における審議が必要であったものが、評価検討会における評価と産業構造審議会による審議は3年に1度行い、その間に終了したものや中間評価の時期に至ったプロジェクトをまとめて行うようになった。個々のプロジェクトについては、毎年モニタリングを行っている。このことによって評価の負担が軽減されている。

さらに、プログラム評価（技術に関する施策の評価）では、主にプログラム全体の目的・政策的な位置づけの妥当性や施策の構造・目的実現の見通し、総合評価を行うことに重点化され、その結果評価者の負担も軽減されている。

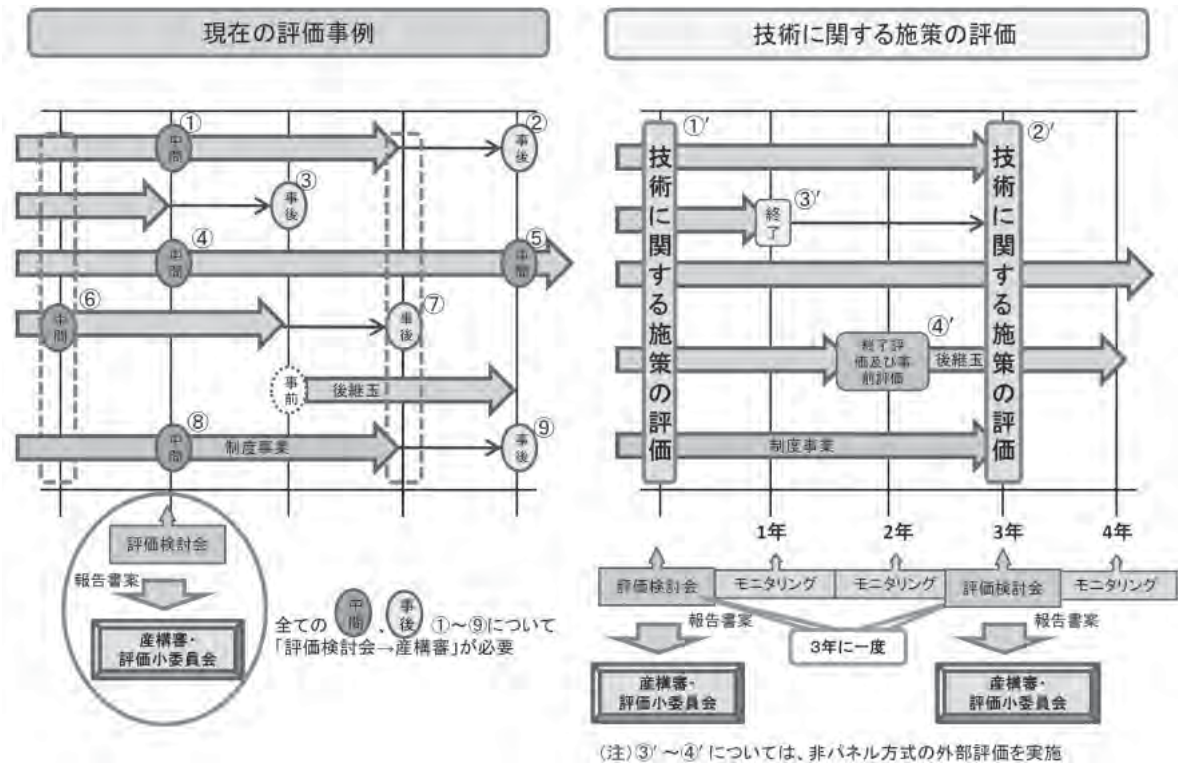
政策目標の達成状況は、プロジェクトレベルではなく、プログラムレベルでの評価に重点がおかれていることがわかる。

(7) 経済産業省ウェブサイト「産業技術政策全般」

<http://www.meti.go.jp/policy/kenkyu_kaihatu/program/program%28toppage%29.htm> [last accessed: 2012/2/7]

(8) 経済産業省産業技術環境局技術評価室「経済産業省技術評価指針に基づく標準的評価項目・評価基準（平成23年7月）」
<http://www.meti.go.jp/policy/tech_evaluation/b00/01/koumoku1107.pdf> [last accessed: 2012/2/14]

図3 これまでの経済産業省のプロジェクト評価と技術に関する施策の評価導入後の比較



(出典) 第25回産業構造審議会評価小委員会資料9 (平成21年1月28日) より抜粋

2 米国

(1) エネルギー省 (Department of Energy: DOE)

DOEの科学局 (Office of Science) には、6つの中核となる「Program Office (室)」が存在しており⁽⁹⁾、その1つが生物環境科学室 (The Office of Biological and Environmental Research: BER) である。生物環境科学室には、生物系環境科学部 (Biological Systems Science Division: BSSD) 及び気候環境科学部 (Climate and Environmental Sciences Division: CESD) の2つの部門 (Division) があり、各部門には部長 (Division Director) がいる。各部門には、地域・地球気候モデリング・プログラム (Regional and Global Climate Modeling: RGCM) やゲノム科学プログラム (Genomic Science Program) などのサブ・プログラムがあり、それぞれにプログラムマネージャーが配置されている。さらにそれぞれ複数の科学研究プロジェクトを抱える階層構造となっている。

DOEにおけるプログラムとは、科学的な研究目標を達成するために設計された一連の活動を指す。各プログラムでは研究目的としてのミッションが定義されると同時に、研究プログラムの重要性として、政策的な意義が強調されている⁽¹⁰⁾。

プログラム内の研究プロジェクトについては3年ごとに研究課題を募集し、研究者から応募された研究課題を競争によって選考するプロセスとなっている。従って、長期的な視点で設計

(9) U.S. DOE Office of Science website, "Programs" <<http://science.energy.gov/programs/>> [last accessed: 2012/2/11]

(10) たとえば, "Regional & Global Climate Modeling (RGCM) Program"

<<http://science.energy.gov/ber/research/cesd/regional-and-global-modeling/>> [last accessed: 2012/2/11]

されたプログラムの下、それぞれの研究プロジェクトは入れ替えがある構造となっている。

プログラム評価は、科学局長官 (Director) の諮問により、生物環境科学室諮問委員会 (Biological and Environmental Research Advisory Committee: BERAC) が行っている。

生物環境科学室諮問委員会の役割は、各プログラムの定期的なレビューを実施することだけでなく、科学的な側面と政策をより効果的に合致させるために長期計画、戦略、ファンディングに対しても答申を行っている。

(2) 環境保護庁 (Environmental Protection Agency: EPA)

EPAの研究開発局 (Office of Research and Development: ORD) では、EPAあるいは他の政府機関のミッションに関連性のある環境分野の主な課題ごとにプログラムが設けられている。

地球変動研究プログラム (Global Change Research Program: GCRP) は、連邦議会が1990年に成立させた「地球変動研究法 (Global Change Research Act)」によって設置された省庁横断型の研究プログラムであり、EPAは同プログラムを構成するメンバー省庁の一つという位置づけとなっている⁽¹¹⁾。地球変動研究プログラムにはビジョンとミッションが定義されており、気候変動等の影響を明らかにし、意思決定に役立てることが挙げられている。

プログラムを統括するのは、全米プログラムディレクター (National Program Director) であり、あらゆるEPAの研究所やセンターで実施する研究とその優先順位の決定について、提携している外部研究機関との調整も含めて責任を持つ。

地球変動研究プログラムについては、年次評価と2~3年に一度のプログラム評価の2つが行われている。研究の優先順位の見直しを行う年次評価では、プログラムの成果等をEPAが報告書としてまとめ、行政管理予算局 (Office of Management and Budget: OMB) がプログラムの評価を行う。プログラム評価では、ノーベル賞受賞者など、世界トップクラスの研究者によって構成された科学顧問委員会 (Board of Scientific Counselors: BOSC) がパネル形式によって、科学的な成果・側面を中心に評価する。評価が終了してBOSCが報告書を発表した後、全米プログラムディレクターは3ヶ月以内にBOSCとのミーティングを開き、評価に対する意見や講じる対策について提示する必要がある、さらにOMBも対策の実施状況を予算編成の参考としている⁽¹²⁾。

(3) 国立衛生研究所 (National Institutes of Health: NIH)

NIHは27の研究所から構成されており、自ら研究を実施すると共にファンディングを行っている。議会は基本的にこれら研究所単位で予算を決定しており、NIHのDirector自身には各研究所への予算配分の権限はない⁽¹³⁾。

NIHは大規模な組織であり、プログラムの編成も一様ではない。一般的に「プログラム」は、特定の分野に対する保健社会福祉省が掲げる特定の分野の目標に基づいている。プログラムに内包される研究プロジェクトでは、プログラム目標に対応した研究を実施する大学に対するグラントなど、より規模や視点が絞られたものを対象としている。プログラムベースで資金を提

(11) “Program Overview” <<http://www.globalchange.gov/about/overview>> [last accessed: 2012/2/11]

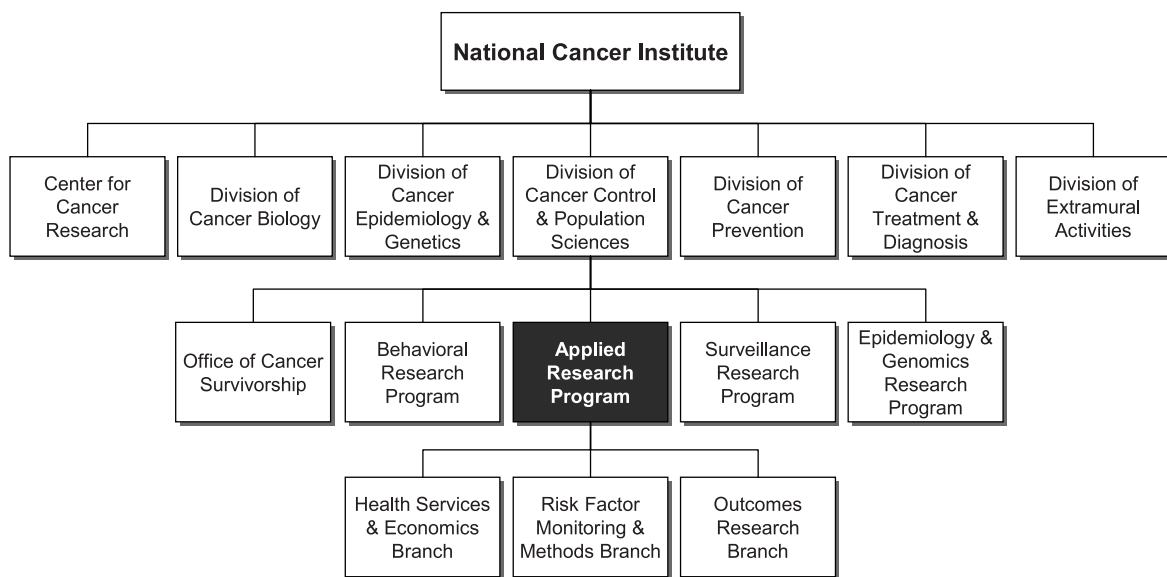
(12) 文部科学省「研究プログラムの構成の考え方とその評価体系、評価手法及び評価者の視点を通じた研究コミュニティの活性化に関する調査・分析報告書 (評価者の視点を通じた研究コミュニティの活性化に関する調査・分析編)」2010、75頁。

(13) 曾根智史「米国National Institutes of Health (NIH) の研究評価制度」『保健医療科学』53巻4号, 2004, 263頁。

供した上で、そのプログラムの中にあるプロジェクトへ分配する仕組みとなっている。⁽¹⁴⁾

国立癌研究センター（National Cancer Institute: NCI）の場合、組織は部門（Division）から構成されているが、部門に内包されるプログラムもあれば、複数部門を横断するプログラムも存在している。一般に、1つのプログラムは相互に関連性がある複数のプロジェクトから構成される。例えば、応用研究プログラム（Applied Research Program）について見ると、NCIの中の癌制御・人口学部門（Division of Cancer Control & Population Sciences）の下にある複数のプログラムの1つとして位置付けられており、さらにプログラムはブランチに分かれている。プログラムの中の優先順位付けや資源配分はプログラムに設けられたThe Office of the Associate Directorが行う。

図4 NCIにおける応用研究プログラムの位置付け



（出典）米国NIH国立癌研究センターウェブサイト<http://appliedresearch.cancer.gov/about/nci_chart.html> [last accessed: 2012/2/11] を基に筆者作成

プログラム評価には、ニーズ評価（Needs Assessment）、実現可能性研究（Feasibility Study）、プロセス評価（Process Evaluation）、アウトカム評価（Outcome Evaluation）の4種類があり、目的に即して選択されている⁽¹⁵⁾。

3 イギリス

(1) 研究会議（Research Councils: RCs）

イギリスでは領域別に7つの研究会議が設けられ、資金配分や傘下の機関による研究の実施を行っている。

(14) 文部科学省「研究プログラムの構成の考え方とその評価体系、評価手法及び評価者の視点を通じた研究コミュニティの活性化に関する調査・分析報告書（研究プログラムの構成の考え方とその評価体系、評価手法に関する調査・分析編）」2010, 154頁。
 (15) 文部科学省『「海外政府系研究開発機関における研究開発評価システムに関する調査・分析」調査報告書』2011, 48頁。

その1つであるバイオテクノロジー・生物科学研究会議（Biotechnology and Biological Sciences Research Council: BBSRC）におけるプログラムは、長期的な政策目標を実現するために設定されるものである。BBSRCにおけるプログラムは、一般的に基礎研究の分野に適用される。

イギリス大蔵省は評価のガイドライン⁽¹⁶⁾を策定しており、プログラム（programme）とは、関連するプロジェクトのグループと定義している⁽¹⁷⁾。BBSRCではこのガイドラインに準拠した評価フレームワーク⁽¹⁸⁾によってプログラム評価を行っており、具体的には、科学的な質、経済社会的な影響を評価することによって以下を達成しようとしている。

- BBSRCが配分した研究資金について利害関係者に説明責任を果たす
- 将来の研究資金配分について意思決定を行う
- BBSRCの研究方針・実践をより良いものにしていく

他の領域の研究会議である工学・物理科学研究会議（Engineering and Physical Sciences Research Council: EPSRC）においても、特定のプロジェクトの詳細や個々の成果よりも投資のポートフォリオに焦点を当て、プロジェクトレベルの成功や失敗よりもプログラムを評価している。例えば、プログラム評価として2002年度から始まったテーマ・デイ（Theme Day）では、助成を受けている研究者を集めて、あらかじめ研究についてのデータを集め、専門家パネルに対するプレゼンテーション、質疑応答を経て、後日報告書を作成している⁽¹⁹⁾。

III まとめ

国による研究開発は、上位である政策から、個別の研究開発課題まで階層的にマネジメントされている。その中でも、政策と研究開発課題・プロジェクトの中間階層に相当するプログラムレベルのマネジメントが重要視されてきている。こうしたマネジメントに対応してプログラムレベルでの研究開発評価が必要である。プログラムレベルでのマネジメント・評価が重視されるのは、個別の研究開発課題をいくらマネジメント・評価しても、政策の実現に結びつけるには限界があることが認識されているためである。

各国の先行事例では、政策目標に対応したプログラムを設定して評価し、プログラムの中のプロジェクトの資源配分はプログラムに委ねるマネジメントが見られる。

たとえば、米国のDOEでは、科学的な研究目標を達成するために設計された一連の活動をプログラムとして定義しており、プログラムレベルのマネジメント・評価を行うことによって、各階層の研究開発を調整し、納税者への成果還元と、研究者の研究目標達成を同時に目指している。NIHでもプログラムベースで資金を提供した上で、その中にあるプロジェクトへ分配する仕組みにすることで、プロジェクトベースのみを支援して視野が狭くならないようにしている。

イギリスの研究会議におけるプログラムでも、長期的な政策目標を実現するために、プログ

(16) “THE GREEN BOOK Appraisal and Evaluation in Central Government”
<http://www.hm-treasury.gov.uk/d/green_book_complete.pdf> [last accessed: 2012/2/11]

(17) “A programme is defined as a group of related projects”

(18) “Evaluation Framework: Research Programmes and Schemes 2011”
<http://www.bbsrc.ac.uk/web/FILES/Reviews/bbsrc_evaluation_framework.pdf> [last accessed: 2012/2/11]

(19) 文部科学省 『「海外政府系研究開発機関における研究開発評価システムに関する調査・分析」調査報告書』2011, 80頁。

ラム評価が行われている。

我が国でも科学技術振興機構における戦略的創造研究推進事業（CREST）で目的を持ったプロジェクト群を戦略目標達成型の大きな研究開発プログラムとして、研究リーダーである研究総括が遂行する例が見られるが、プログラムレベルでのマネジメント、プログラムレベルでの研究開発評価は試行錯誤が行われている段階である。