

<資料>

日・米・英の保健医療に関連する競争的研究資金における研究課題評価の比較

武村真治

国立保健医療科学院健康危機管理研究部

Comparison of research proposal assessment in competitive research funding related to healthcare among Japan, U.S.A., and U.K.

Shinji TAKEMURA

Department of Health Crisis Management, National Institute of Public Health

抄録

目的：アメリカのNational Institutes of Health (NIH), イギリスのMedical Research Council (MRC), わが国の厚生労働科学研究費補助金(厚労科研)における研究開発課題の評価システムの実態を把握・比較し, 諸外国の取り組みのわが国への適用可能性を検討する。

方法：NIH, MRC, 厚生労働省のホームページで公開されている研究開発課題評価に関する資料を収集し, 運営管理体制, 評価手順, 評価基準等に関して国際比較を行い, それぞれの特徴を記述した。

結果：NIH, MRCでは申請書の評価が「二段階」で実施されていた。NIHでは, 一次評価は学問領域・専門分野別の科学審査グループにおける査読及び評価会議によって, 二次評価はNIHの各研究所に設置される諮問委員会によって, それぞれ実施されていた。またMRCでは, 一次評価は専門の審査員による査読によって, 二次評価は学問分野別の研究委員会, または特定の研究費助成スキーム(研究事業)に設置されるパネルによって実施されていた。また評価作業の合理化を図るために, 申請書の「トリアージ」が行われ, 査読において評価が低い申請書は議論や二次評価の対象から除外されていた。

厚労科研の評価基準(重要性, 発展性, 独創性・新規性, 実現性・効率性, 研究者の資質・施設的能力, 政策等への活用, 行政的緊急性, 効率的・効果的な運営の確保), NIHの主要評価基準(重要性, 研究者, 革新性, アプローチ, 研究環境), MRCの中核の評価基準(重要性, 科学的潜在能力(研究環境と研究者, 研究計画), 資源の正当性)はほとんど同じ領域をカバーしていたが, NIH, MRCでは各評価基準における評価の視点がかつ詳細に設定されていた。

結論：厚労科研に二段階評価の仕組みを導入し, 一次評価におけるトリアージで学術的価値の高い申請課題を選択し, 二次評価で行政的価値の高い申請課題を選択することによって, 学術的側面と行政的側面ともに評価の高い研究開発課題を採択することが可能になること, 評価の信頼性や妥当性の向上, 厚労科研の目的である「行政施策への貢献」を重視した評価のために, 各評価基準に対する下位の基準や具体的な評価の視点を明示する必要があること, が示唆された。

キーワード：研究評価, 評価基準, 厚生労働科学研究費補助金, 研究費配分機関 (Funding Agency: FA), 医学研究

連絡先：武村真治

〒351-0197 埼玉県和光市南2-3-6

2-3-6, Minami, Wako-shi, Saitama, 351-0197, Japan.

Tel: 048-458-6166

Fax: 048-468-7983

E-mail: takeshin@niph.go.jp

[平成25年12月16日受理]

Abstract

Objectives: To compare the current status of the research proposal assessment systems of the National Institutes of Health (NIH) in the United States, the Medical Research Council (MRC) in the United Kingdom, and the Health and Labour Sciences Research Grants (HLSRG) in Japan, and to examine the possibility of applying the approaches followed in other countries to Japan.

Methods: We collected documents related to research proposal assessments posted on the websites of the NIH, MRC, and Ministry of Health, Labour and Welfare, Japan. We compared such factors as operational/management structures, assessment procedures, and assessment criteria among the three countries, and described the characteristics of each.

Results: Application forms were assessed in two stages at the NIH and MRC. Peer review and assessment meetings of the Scientific Review Group from each academic discipline and specialized field were conducted during the initial NIH assessment. National Advisory Councils/Boards established in each research institute of the NIH performed the second assessment. At the MRC, a peer review of expert reviewers conducted the initial assessment and a research board from each academic discipline or a panel established by a specified research grant scheme conducted the second. To streamline the assessment process, a “triage” was performed on application forms; those that received low peer review assessments were excluded from the second assessment.

The same areas were mostly covered in the HLSRG assessment criteria (significance, expansibility, originality/novelty, feasibility/efficiency, researcher qualifications and institutional capacity, applicability to policy, administrative urgency, and securing efficient and effective operations), the scored review criteria of the NIH (significance, investigator(s), innovation, approach, and environment), and the core criteria of the MRC (importance, scientific potential [environment and people, research plans], and justification for resources). However, at the NIH and MRC, assessment perspectives on each criterion were more specifically addressed and described in detail.

Conclusions: The results suggested that it might be possible to adopt research proposals that received high academic and administrative assessments by introducing two-stage assessments at the HLSRG. In addition, research proposals with high academic value should be selected by “triage” during the initial assessment and those with high administrative value during the second. Furthermore, it might be necessary to present sub-criteria for each assessment criterion and specific assessment perspectives to ensure that they focus on enhancing reliability and validity and contribute to policy-making, which is the aim of HLSRG.

keywords: research evaluation, assessment criteria, Health and Labour Sciences Research Grants, funding agency (FA), medical research

(accepted for publication, 16th December 2013)

I. はじめに

国民の健康水準を向上させるためには、効果的な保健医療福祉施策を展開する上で拠り所となる「研究開発」を積極的に推進していく必要がある。研究開発に対しては様々な資金が投資されているが、その中でも「競争的資金制度」については毎年拡充が図られており [1]、一層の充実が求められている [2]。競争的資金制度は、研究者等から提案された複数の課題から優れたものを採択し、研究資金を配分する、というシステムで運用されているが、研究資金の効率的な配分のためには、より多くの研究成果が期待される課題を選択するための「研究開発課題の評価」 [3] を適正に実施する必要がある。

わが国の保健医療に関連する競争的資金制度の代表的なものとして「厚生労働科学研究費補助金」(以下、厚

労科研)が挙げられる。厚労科研は1951年から開始され [1]、国民の保健医療、福祉等に関する行政施策に資する研究開発課題の評価の仕組みを構築してきた [4] が、今後もさらに発展させるためには諸外国の経験や先進事例を活用することが有用であると考えられる。

諸外国における保健医療関連の研究費配分機関 (Funding Agency: FA) としては、アメリカのNational Institutes of Health (NIH)、イギリスのMedical Research Council (MRC) が代表的である。NIHは1930年 (前身は1887年) に、MRCは1913年に、医学研究を総合的に推進することを目的として設立された組織である。NIHは複数の研究所・センター (Institute and Center) で構成され、各研究所が自ら研究を実施する (intramural research) とともに、大学等の外部の研究機関に研究費を配分して研究を実施してもらう (extramural research) 役割を担っている。MRCも同様に、intramural research

とextramural researchを実施している。

わが国の厚労科研を含む競争的資金制度はNIH, MRCにおけるextramural researchの研究費配分システムに相当し、それらの研究開発課題の評価の実態は報告されている [5, 6] が、最新の取り組みの把握やわが国との厳密な比較が行われていないため、わが国への適用可能性が十分に検討されていないのが現状である。

そこで本研究は、アメリカのNIH, イギリスのMRC, わが国の厚労科研における研究開発課題評価を比較し、諸外国の取り組みを踏まえて、わが国における効果的な研究開発課題の評価システムを構築するための具体的な方策を検討することを目的とした。

II. 方法

ホームページで公開されている研究開発課題の評価に関連するNIH [7], MRC [8, 9], 厚労科研 [4, 10] の資料を収集し、その運営管理体制, 評価手順, 評価基準等に関する情報を抽出し、諸外国との比較においてわが国の長所, 短所を抽出し、諸外国の取り組みのわが国への適用可能性を検討した。

III. 結果

1. 競争的資金制度の概要

NIHは、「患者を助けること」、「医学を前進させること」を目標として、基礎研究, 臨床研究, トランスレーショナルリサーチを含む医学研究を総合的に実施する組織である。24の研究所 (Institute) とセンター (Center) で構成され、各研究所・センターがintramural researchとextramural researchの研究費を運営管理している。extramural researchの予算は2012年度で約250億ドルである。1課題に対する研究費の規模に関しては、NIHの一般的な補助金であるResearch Project Grant Program (R01) で、1年間で最大25万ドル、最長5年間の助成がなされる。R01の採択課題数は2012年度で5,340件、採択率は18%である。評価委員会 (スタディセクション) の委員数は20~30名である。

MRCは、人々の健康を改善し、それによって幅広い利益、特に経済的な利益 (生産性の向上と所得・税収の増加, 労働時間の増加, 教育への投資の増加とそれによる生産性の向上, 長寿による貯蓄の増加とさらなる投資の増加) を得ることを目的として、基礎研究を中心として、臨床研究, 公衆衛生研究などの幅広い医学研究を実施する組織である。extramural researchの予算は2011年度で約3.10億ポンドである。1課題に対する研究費の規模に関しては、MRCの一般的な補助金であるResearch grantで、1年間で最大20万ポンド、最長5年間の助成がなされる。Research grantの採択課題数は2012年度で280件、採択率は26%である。評価委員会 (研究委員会) の委員数は25名程度である。

厚生労働科学研究費補助金は、研究の結果およびエビデンスを厚生労働行政における各種の施策に反映させ、また施策の成果をエビデンスとして把握することにより、国民の健康・安全確保を推進することを目的として、国民の保健医療, 福祉, 生活衛生, 労働安全衛生等の様々な課題を解決する「目的指向型の研究課題」を推進している。研究費の予算は2012年度で約440億円である。1課題に対する研究費の規模は研究事業によって異なるが、2010年度で平均約2,900万円で、最長3年間の補助がなされる。採択課題数は2010年度で759件、採択率は34.0%である。評価委員会の委員数は10~15名が標準とされる。

2. 研究開発課題の公募・受付

アメリカ, イギリス, 日本のいずれも、公募要項 (NIHではFunding Opportunity Announcement (FOA), MRCではcall for proposal) を公布して、幅広く研究開発課題を募集している。NIH, MRCともに、研究者自身が企画した研究テーマを募集することが多いが、NIHの研究所, MRCの研究費助成スキーム (研究事業) が特定のテーマを設定して公募する場合もある。また公募時期は随時、あるいはNIHの各研究所やMRCの各研究費助成スキームによって異なる。それに対して厚労科研では、ほとんどの研究テーマは各研究事業で設定され、公募時期も毎年10月~11月と研究事業で共通である。

申請書の受付に関しては、NIHでは、extramural researchの研究開発課題評価をNIH横断的に運営管理するCenter for Scientific Review (CSR) が一括して実施し、MRCでは学問分野別に設置されている研究委員会 (Research Board), または特定の研究費助成スキームに設置されるパネルが実施する。厚労科研では各研究事業の所管課室が申請書 (研究計画書) を受け付けている。

3. 研究開発課題の評価

1) 評価の「段階」

NIHもMRCも、申請書の評価を「二段階」で実施している。

NIHでは、一次評価 (Initial Level of Review) は、学問領域・専門分野別の学識経験者で構成される複数の科学審査グループ (Scientific Review Group: SRG) によって実施される。CSRに提出された申請書はその内容に適合したSRGに送付され、まず申請書に対してSRGのメンバー (最低3名) が割り当てられ、申請書の「査読」が行われる。そしてSRGの評価会議において査読結果 (採点, コメント) が公表され、それに関して議論を行った後にメンバー全員が採点を行い、最終的な評価点数を確定する。また二次評価 (Second Level of Review) は各研究所に設置される諮問委員会 (National Advisory Council or Board) によって実施される。そこでは、一次評価の結果を踏まえた上で、一次評価における科学的・技術的な側面だけでなく、各研究所の研究開発方針

や優先度等との関連性についても評価を行い、採択課題が決定される。

MRCでは、一次評価は専門の審査員による査読（ピア・レビュー）によって行われ、各申請書に対して採点とコメントの記述が行われる。二次評価は研究委員会・パネルによって実施され、一次評価の点数とコメントを参考にして、MRCの研究開発戦略や国の研究開発方針への適合性の観点からも評価を行い、研究委員会・パネル全体として採点を行い、最終的な評価点数を確定し、採択課題が決定される。

厚労科研では、各研究事業に設置される事前評価委員会によって評価が行われる。まず各申請書（研究計画書）に対して事前評価委員が割り当てられ、査読（採点、コメントの記述）が行われる（書面評価）。査読結果は事務局によってとりまとめが行われ、事前評価委員会において提示される。そして委員会の議論を経た後に採択課題が決定される。

2) 申請書の「トリアージ」

NIHもMRCも、提出される膨大な数の申請書に対する評価作業の合理化を図るために、申請書のトリアージが行われる。NIHでは、SRGのメンバーによる査読において評価の高い申請書のみが評価会議で議論され、評価の低い申請書は議論されることも最終的な評価点数がつけられることもない。MRCも、一次評価での評価点数が低く、採択される可能性の低い申請書は二次評価の対象から除外される。一方、厚労科研では、事前評価委員会において申請書のトリアージが評価手順として明確には規定されていない。ただし、必要に応じて評価小委員会を設置できることが規定されており、特に評価対象となる申請書が多い研究事業では設置されている場合があ

る。その際に、評価小委員会の評価結果がトリアージに利用されている可能性がある。

3) 評価の基準

NIHの一次評価では、主要評価基準、補足的評価基準、補足的評価検討項目が設定され、主要評価基準ごとに採点とコメントの記述、補足的評価基準ごとにコメントの記述が行われる。そしてそれらを参考にして、全体的影響力（overall impact）の総合評価が行われる。主要評価基準として、重要性（Significance）、研究者（Investigator (s)）、革新性（Innovation）、アプローチ（Approach）、研究環境（Environment）の5領域が設定されている。

MRCでは、一次評価、二次評価に共通する中核的評価基準として、重要性（importance）、科学的潜在能力（scientific potential）、（要求される）資源の正当性（justification for resources）の3領域が設定されている。また科学的潜在能力は「研究環境と研究者」、「研究計画」の2つに細分化されている。なお二次評価では、各領域について、一次評価に追加する評価基準が設定されている。

厚労科研の事前評価では、「(1) 専門的・学術的観点からの評価に当たり考慮すべき事項」として、厚生労働科学分野における重要性及び発展性、研究の独創性・新規性、研究目標の実現性・効率性、研究者の資質、施設的能力、「(2) 行政的な観点からの評価に当たり考慮すべき事項」として、政策等への活用（公的研究としての意義）、行政的緊急性、「(3) 効率的・効果的な運営の確保の観点からの評価に当たり考慮すべき事項」が設定されている。

表1～6に、日本（厚労科研の事前評価）、アメリカ（NIHの一次評価）、イギリス（MRCの一次評価及び二

表1 厚生労働科学研究費補助金(日本), NIH(アメリカ), MRC(イギリス)の研究開発課題の評価基準の比較(研究の重要性, 発展性)

厚労科研	NIHの一次評価	MRCの一次評価	MRCの二次評価
(1) ア. 研究の厚生労働科学分野における重要性 ・厚生労働科学分野に関して有用と考えられる研究であるか	○重要性 ・当該研究分野における重要な問題や、進歩に対する重大な障害に取り組んでいるか？ ・研究目的が達成された場合、科学的知識、技術的能力、臨床での実践はいかに改善されるか？ ・目標達成に成功すれば、当該研究分野を推進する理論的概念、方法、技術、治療、サービス、予防介入はどのように変革されるか？	○重要性 ・この分野の研究は必要とされているか？ ・取り上げられている問題や未解明点に関して、探究するだけの十分な医学上、科学上の論理的根拠はあるか？ 成功すれば新たに重要な知見を得られる見込みはあるか？ ・人間の健康の改善に関して、得られると見込まれる最終的な便益が現実的に示されているか？ ・疾病負荷の軽減に関して、直接的、間接的にどの程度貢献するか？ ・現時点で本研究を行うことはどれほど重要か？ ・他で進行している類似の研究や相補的な研究はないか？ 競争力はあるか？	○重要性 ・例えば、新たな臨床上の進歩を実践に導入できる機会といった、重要な「チャンス」はあるか？ ・研究結果の再現性や他のシステムへの拡張可能性に関する論拠はどの程度説得力があるか？
(1) イ. 研究の厚生労働科学分野における発展性 ①研究成果が厚生労働科学分野の振興・発展に役立つか ②臨床研究の場合は、いわゆる臨床研究登録がなされる予定か		○科学的潜在能力 (研究計画) ・試験的研究や原理証明を扱う研究の場合、本研究はどのように発展することになるのか？ またそれはどの程度実現可能か？	○科学的潜在能力 (研究計画) ・応用研究、基礎研究に関わらず、研究成果を他の領域や分野における実践の改善につなげるための現実的なアプローチは設定されているか？

(注) ○は主要または中核的評価基準、※はそれ以外の評価基準を示す。

次評価)の研究開発課題評価の基準の比較表を示した。なお、MRCの二次評価は一次評価に追加する基準のみを示した。厚労科研、NIH、MRCを比較すると、評価基準の範囲の設定は若干異なる部分もあるが、全体としてカバーする評価の領域はほとんど同じであった。しかし厚労科研と比較して、NIH、MRCは各評価基準における評価の視点がより詳細に設定されていた。

個別の評価基準で比較すると、研究者と研究環境に関して、厚労科研では「(1)オ.研究者の資質、施設の能力」、MRCでは「科学的潜在能力(研究環境と研究者)」と、一つの評価基準になっていたが、NIHでは「研究者」と「研究環境」に明確に分離されていた(表4)。

「政策等への活用」に関しては、NIHの一次評価では評価基準として設定されていなかった。またMRCでは、

研究成果の商業的利用や知的財産などの経済的側面が強調されていた(表5)。

「効率性・効果的な運営」に関しては、NIHの一次評価では主要評価基準としては設定されていなかった。MRCと厚労科研を比較すると、MRCでは研究に投入される時間、スタッフ等の資源(費用)自体の効率性が評価されるのに対して、厚労科研では民間研究などの他の研究との代替可能性や連携可能性の観点からの効率性が評価されていた(表6)。

厚労科研のみで設定されていた評価基準として、(1)のイの②、オの②のように、臨床研究に係る評価基準(臨床研究登録、疫学・生物統計学の専門家の関与)(表1、表4)、(2)のイの「行政的緊急性」(表6)が挙げられた。

表2 厚生労働科学研究費補助金(日本)、NIH(アメリカ)、MRC(イギリス)の研究開発課題の評価基準の比較(研究の独創性・新規性)

厚労科研	NIHの一次評価	MRCの一次評価	MRCの二次評価
(1)ウ. 研究の独創性・新規性 ・研究内容が独創性・新規性を有しているか	○革新性 ・斬新な理論的概念、アプローチや方法論、装置、介入などを利用して、現在の研究や臨床での実践のパラダイムシフトに挑戦し、それを探究しているか? ・その概念、アプローチ、方法論、装置、介入などは、特定の研究分野において斬新であるか、あるいは広い意味において斬新であるか? ・理論的概念、アプローチ、方法論、装置、介入に関して、高度化、改善、新規活用法が提案されているか?	○科学的潜在能力(研究計画) ・研究計画はどの程度革新的か? ・研究方法や研究デザインは当該研究分野の最高のものと比較して競争力があるか?	○科学的潜在能力(研究計画) ・研究計画全般や個々の研究方法はどの程度革新的か、あるいはそれは試行済み、試験済みのアプローチであるか? そのアプローチが失敗する危険性が高いと判断される場合、成功した場合に得られる便益との釣り合いが取れているか?

(注) ○は主要または中核的評価基準、※はそれ以外の評価基準を示す。

表3 厚生労働科学研究費補助金(日本)、NIH(アメリカ)、MRC(イギリス)の研究開発課題の評価基準の比較(研究の実現性・効率性)

厚労科研	NIHの一次評価	MRCの一次評価	MRCの二次評価
(1)エ. 研究目標の実現性・効率性 ①研究期間の各年度毎の目標が明確か	○アプローチ ・全体の戦略、方法論および分析は、研究課題が目指す目標を達成するにあたって合理的かつ適切か?		
(1)エ. 研究目標の実現性・効率性 ②実現可能な研究であるか	○アプローチ ・目標達成に関わる潜在的な問題、代替戦略、ベンチマークは提示されているか? ・研究開発課題が初期の発展段階にある場合、実現可能性を確立するための戦略が示されているか、また特に研究のリスクの側面は管理されているか?	○科学的潜在能力(研究計画) ・研究計画は研究の目的や研究資源等からみて現実的か? ・科学面、技術面、組織面などにおける主要な問題が同定されているか、またそれらへの対応は十分か?	○科学的潜在能力(研究計画) ・査読者から提起された技術的コメントを考慮して、提案されたアプローチはどの程度実現可能か? またそれは当該の問題に取り組む上で最も効果的な方法となる見込みはあるか? ・価値を付加するための研究の組織化と連携の促進に関する計画は、質や生産性のより高い科学の実現に向けて、現実的かつ見込みのあるものであるか?
	○アプローチ ・臨床研究が含まれる場合、1)研究のリスクから被験者を保護する計画、2)少数民族、男女両性、未成年を被験者に含める計画は、科学的目標と研究戦略の観点から正当化できるか?		

(注) ○は主要または中核的評価基準、※はそれ以外の評価基準を示す。

表4 厚生労働科学研究費補助金(日本), NIH(アメリカ), MRC(イギリス)の研究開発課題の評価基準の比較(研究者の資質, 施設の能力)

厚労科研	NIHの一次評価	MRCの一次評価	MRCの二次評価
(1) オ. 研究者の資質, 施設の能力 ①研究業績や研究者の構成, 施設の設備等の観点から, 遂行可能な研究であるか ②臨床研究の場合は, 疫学・生物統計学の専門家が関与しているか	○研究者 ・主任研究者, 共同研究者などの研究者は当該研究課題への適性があるか? ・経験の浅い研究者, 新人の研究者の場合, 適切な経験を持ち, 適切な訓練を受けているか? ・確立された研究者の場合, 当該研究分野の発展に寄与している進行中の業績が提示されているか? ・研究課題が共同研究または複数の主任研究者による研究の場合, 研究者たちは相補的かつ統合的な専門性をもっているか? ・リーダーシップのアプローチ, ガバナンス, 組織構成は研究開発課題に適しているか? ○研究環境 ・研究開発課題が実施される科学的環境は成功の可能性を高めることに貢献するか? ・研究機関の支援, 設備など, 研究者が利用できる物理的資源は研究開発課題にとって十分か? ・研究開発課題は, その科学的環境に固有の特徴, 研究対象集団, 共同研究体制から便益を享受できるか?	○科学的潜在能力(研究環境と研究者) ・研究者または研究グループは当該研究分野において高い質の実績を確立しているか? ・申請者は本研究を実施できる体制を独自に整えているか? ・申請者にとってなじみのない分野の研究に取り組む申請書である場合, または初めての資金供与の申請書である場合, その研究を推進するための確固たる基盤はあるか? ・本研究は申請者が進めている他の関連する研究にどの程度適合しているか? ・所属研究機関は本研究を支援する意志を明示しているか?	○科学的潜在能力(研究環境と研究者) ・これまでに実施してきた研究は, 効果的に活用されてきたか, ヘルスケアに実質的な効果をもたらしたか?

(注) ○は主要または中核的評価基準, ※はそれ以外の評価基準を示す。

表5 厚生労働科学研究費補助金(日本), NIH(アメリカ), MRC(イギリス)の研究開発課題の評価基準の比較(政策等への活用)

厚労科研	NIHの一次評価	MRCの一次評価	MRCの二次評価
(2) ア. 政策等への活用(公的研究としての意義) ・施策への直接反映の可能性あるいは, 政策形成の過程などにおける参考として間接的に活用される可能性 ・間接的な波及効果などが期待できるか ・これら政策等への活用がわかりやすく具体的かつ明確に示されているか ・社会的・経済的効果が高い研究であるか		※商業的利用 ・商業的利用が可能な研究成果を生み出す見込みはあるか? ・所属研究機関は, 本研究による知的財産の商業的開発を推進する体制を十分に整えているか? またその場合において提案されている仕組みは適切か?	○重要性 ・ヘルスサービスや健康関連業界に対して, 直接的に重要性のある成果が生み出される見込みはあるか? ・MRCやその他の主要な機関(例えば, Department of Health, Department for International Developmentなど)によって特定されたニーズや戦略に効果的に取り組んでいるか? ○科学的潜在能力(研究計画) ・本研究で得られる知的財産の商業的開発に向けた現実的なアプローチは設定されているか?

(注) ○は主要または中核的評価基準, ※はそれ以外の評価基準を示す。

表6 厚生労働科学研究費補助金(日本), NIH(アメリカ), MRC(イギリス)の研究開発課題の評価基準の比較(行政的緊急性, 効率的・効果的な運営)

厚労科研	NIHの一次評価	MRCの一次評価	MRCの二次評価
(2) イ. 行政的緊急性			
(3) 効率的・効果的な運営の確保の観点からの評価に当たり考慮すべき事項 ・研究が効果的・効率的に実施(計画)されているか ・他の民間研究などにより代替できるものではないか ・研究の実施に当たり, 他の公的研究・民間研究などとの連携や活用が十分に図られているか(他の公的研究・民間研究や過去の成果などの活用, 共同研究とすることなどによる研究効率や効果の向上の余地がないか)	※予算と研究期間(補足的評価検討項目) ・査読者は, 要求されている予算や研究期間が, 提案されている研究に照らして十分に正当かつ合理的かどうかを考察する。	○研究資源の正当性…要求されている資金はその研究を遂行する上で不可欠であるか? また資金の規模は, 重要性及び科学的潜在能力からみて妥当か? ・申請者は必要とする全ての資源を特定しなければならず, また審査員は全てについて評価を行う(資産の費用と間接費用を除く). ・主任研究者, 共同研究者が研究に投入する時間は, 彼らの研究への関与の程度に合致しているか, 研究を良好に管理する上で必要または十分か, 彼らが対応できる現実的な時間の見積りとなっているか? ・その他の研究スタッフの人数は記載されている研究内容に相応しいか? ・要求されている資源に対する正当な根拠が明示されているか? ・要求されている資源からみて, 費用に見合う価値(value for money)を示しているか?	○研究資源の正当性 ・スタッフや研究資源全般の数量に関する請求の理由を明示しているか? その規模は記載されている研究活動にとって十分か? ・提案されている研究活動や要求されている研究資源の規模から期待される便益を考慮して, 費用に見合う価値(value for money)を十分に得られる見込みはあるか?

(注) ○は主要または中核的評価基準, ※はそれ以外の評価基準を示す。

IV. 考察

1. 研究開発課題評価の「段階」の明確化

NIH, MRCでは、申請書の評価が「二段階」で実施されていた。一次評価では「学術的・科学的観点」を中心に、二次評価ではそれに加えて、FAあるいは国の研究開発戦略との適合性といった「政策的観点」を中心に、それぞれ異なる視点での評価が実施されていた。また査読におけるトリアージを経た後に二次評価が実施されていた。トリアージは、評価作業の負担を軽減すると同時に、競争力の高い申請書を排除しない公平な方法であり [11]、二次評価の対象となる申請書の学術的・科学的な質を確保することを可能にすると考えられる。一方、厚労科研では、事前評価委員会において学術的側面と行政的側面の両面からの評価が「一段階」で行われる。そのため、両側面が「平均」された形で総合評価される可能性があり、学術的評価が低い課題でも行政的評価が高ければ、あるいはその逆でも、採択されてしまう可能性がある。

NIH, MRCともに、「学術的観点」で一次評価、「政策的観点」で二次評価を実施しているが、逆の順序であっても最終的な結果は変わらない可能性がある。この順序で評価する理由として、NIH, MRCは医学研究の推進を目的としているため、学術的な質の確保を優先していることが挙げられる。また政策的観点を重視する場合は契約 (Contract) 等の形で実施する枠組みがあるため、研究課題を「公募」する必要性は低いと考えられる。「公募」は、国やFAが想定していないような、つまり政策的観点では捉えられないような独創的、革新的な研究を広く募集することが目的であり、その評価においては学術的観点が優先される必要があると考えられる。

厚労科研では、国民の保健医療、福祉等に関する行政施策に資する研究課題が求められており、NIH, MRCと比較して行政的側面が重視される傾向があると考えられる。しかし行政施策に資する「エビデンス」の確立を目的として研究課題が「公募」される以上、NIH, MRCと同様に、学術的側面でも評価がなされる必要があると考えられる。また、行政的側面を重視するのであれば、「契約」あるいは「指定研究」の形で実施する方が効率的であると考えられる。

したがってわが国においても二段階評価の仕組みを導入し、一次評価におけるトリアージで学術的価値の高い申請課題を選択し、二次評価で行政的価値の高い申請課題を選択することによって、学術的側面と行政的側面ともに評価の高い研究開発課題を採択することが可能になると考えられる。特に二次評価に関しては、学識経験者だけでなく、政策立案者、実務者、関係団体、産業界などのステークホルダーが関与することによって政策・行政への反映が促進される可能性があり [12]、厚労科研の目的である、国民の保健医療、福祉等に関する行政施策

に資する研究開発の促進に貢献できると考えられる。

2. 研究開発課題の評価基準の改善

1) 評価基準の具体化

厚労科研の評価基準 (重要性, 発展性, 独創性・新規性, 実現性・効率性, 研究者の資質・施設的能力, 政策等への活用, 行政的緊急性, 効率的・効果的な運営の確保), NIHの主要評価基準 (重要性, 研究者, 革新性, アプローチ, 研究環境), MRCの中核的評価基準 (重要性, 科学的潜在能力 (研究環境と研究者, 研究計画), 資源の正当性) はほとんど同じ領域をカバーしていたが、評価基準における評価の視点に関しては、厚労科研はNIHやMRCと比較して具体性が低かった。多くの研究 [13-18] で、評価の信頼性、つまり評価者間の採点のばらつきの問題が指摘されている。この問題に対して、1申請課題を評価する評価者数の増加 [13]、評価者の採点パターン (高く、あるいは低く採点する傾向) の調整 [14]、級内相関係数に基づく調整 [15] などの統計学的な調整方法も検討されているが、ばらつきの要因である評価基準に対する評価者の認識の違いを解決するためには評価基準の明確化やガイドラインの設定が必要である [17, 18] とされている。また、多くの評価者はFAから提示される評価のガイダンスが明確でないと認識しており [19]、多くの申請者は評価基準の重みづけの方法に疑問をもっている [20]、という意識調査の結果もある。したがって評価の信頼性や妥当性を向上させる上でも、各評価基準に対する下位の基準や具体的な評価の視点を明示する必要があると考えられる。

2) 評価基準の重点化

評価基準がカバーする領域は厚労科研, NIH, MRCでほとんど同じであったが、重点を置く評価基準や評価の視点に特徴がみられた。例えばMRCでは、中核的評価基準として資源の正当性が設定されていること、研究成果の商業的利用が評価対象となっていることから「経済的側面」が重視されていると考えられる。

一方、厚労科研では、「行政的緊急性」、「政策等への活用 (公的研究としての意義)」が評価されること、民間研究との対比で「効率的・効果的な運営の確保」が求められていることから、「行政的側面」、つまり研究開発課題の行政施策への貢献度が重視されていると考えられる。これは厚労科研の目的であり、重視されるのは当然のことであるが、上述したように評価の視点が具体的に明示されていないため、行政的側面を重視した評価が十分にできていない可能性がある。したがって厚生労働行政に真に必要な研究開発を推進するためにも、行政的側面に関する具体的な評価基準や評価指標を開発する必要がある。

3. 本研究の問題点と今後の課題

本研究は、ウェブ上で公開されている資料を用いて比較分析を行ったが、NIHの二次評価に関する情報が公開

されていなかったため、その実態を把握することができなかった。NIHの二次評価に関しては、研究内容に関する査読や採点は行われず [5, 21]、一次評価が公正で信頼に足るものか [5]、各研究所の使命や方針と合致しているか [21]、倫理上の問題はないか [21]、などの観点で採択の可否が決定される。また、各研究所が配分できる予算額が異なるため、一次評価で同じ点数であったとしても、どの研究所で二次評価が行われるかによって採択の可否が異なる場合がある [21, 22]。このように、各研究所の方針や予算によって採択の可否が異なるため、二次評価における評価基準を具体的に明示できない可能性があると考えられる。またNIH, MRCは、それぞれ様々な研究開発プログラム（研究費助成スキーム、研究事業等）を実施しているが、本研究ではそれらに共通する研究開発課題評価を対象としたため、個別のプログラムの評価に関する分析を行わなかった。今後は現地訪問調査などを実施してより詳細な分析を行う必要がある。

諸外国では、上述した研究開発課題の評価の信頼性 [13-18, 23-26] の他に、評価の公平性 [27-30] に関する実証研究も数多く実施されている。具体的には、臨床研究は基礎研究と比較して採択率や評価点数が低く、公平に評価されていないことが指摘されている [27, 28]。その理由の一つとしてヒトの保護への取り組みが不十分である申請が多いこと [30] が挙げられるが、NIHの評価基準が改訂されたことによって臨床研究に対する評価が改善したという報告 [29] もある。これらは全て評価点数等のデータを用いた実証研究であり、このような研究が評価の信頼性や公平性の向上に寄与していると考えられる。したがって、わが国においても、厚労科研の研究開発課題評価に関するデータ（評価点数、採択率等）を用いた分析を行い、効果的な評価手法の開発に資する実証研究を推進する必要がある。

謝辞

本研究は、厚生労働科学研究費補助金（健康安全・危機管理対策総合研究事業）（課題番号H24-健危-指定-001）の助成によって行われた。

文献

- [1] 浅見真理, 高階恵美子. 我が国の保健医療分野の研究助成. 保健医療科学. 2004;53:252-61.
- [2] 閣議決定. 科学技術基本計画. 2011-8-19. http://www.mext.go.jp/component/a_menu/science/detail/_icsFiles/fieldfile/2011/08/19/1293746_02.pdf (accessed 2013-10-12)
- [3] 内閣総理大臣決定. 国の研究開発評価に関する大綱的指針. 2012-12-6. <http://www8.cao.go.jp/cstp/output/20121206sisin.pdf> (accessed 2013-10-12)
- [4] 厚生労働省大臣官房厚生科学課. 厚生労働省の科学研究開発評価に関する指針. 2010-11-11. <http://www.mhlw.go.jp/bunya/kenkyuu/jigyuu/hyouka3/pdf/01a.pdf> (accessed 2013-4-4)
- [5] 曾根智史. 米国National Institutes of Health (NIH) の研究評価制度. 保健医療科学. 2004; 53: 262-70.
- [6] 緒方裕光, 西村秋生, 伊藤弘人. 諸外国, 国際機関における研究評価制度—英国, OECD, EUおよびWHOにおける研究評価—. 保健医療科学. 2004; 53:276-85.
- [7] U.S. Department of Health and Human Services National Institutes of Health. NIH Grants Policy Statement. 2012-10-1. http://grants.nih.gov/grants/policy/nihgps_2012/nihgps_2012.pdf (accessed 2013-4-4)
- [8] Medical Research Council. Handbook for Applicants and Grant Holders. 2012-11-30. http://www.mrc.ac.uk/consumption/idcplg?IdcService=GET_FILE&dID=39404&dDocName=MRC001873&allowInterrupt=1 (accessed 2013-4-4)
- [9] Medical Research Council. Reviewers Handbook|2011-2012. 2011-5-12. http://www.mrc.ac.uk/consumption/idcplg?IdcService=GET_FILE&dID=38663&dDocName=MRC003184&allowInterrupt=1 (accessed 2013-4-4)
- [10] 厚生労働省大臣官房厚生科学課. 平成25年度厚生労働科学研究費補助金公募要項. 2012-11-9. <http://www.mhlw.go.jp/bunya/kenkyuu/jigyuu/hojokin-koubo-h25/dl/koubo.pdf> (accessed 2013-4-4)
- [11] Vener KJ, Feuer EJ, Gorelic L. A statistical model validating triage for the peer review process: keeping the competitive applications in the review pipeline. FASEB J. 1993;7:1312-9.
- [12] Street J, Baum F, Anderson IP. Is peer review useful in assessing research proposals in Indigenous health? A case study. Health Res Policy Syst. 2009;7:2. doi: 10.1186/1478-4505-7-2.
- [13] Kaplan D, Lacetera N, Kaplan C. Sample size and precision in NIH peer review. PLoS One. 2008;3:e2761. doi: 10.1371/journal.pone.0002761.
- [14] Johnson VE. Statistical analysis of the National Institutes of Health peer review system. Proc Natl Acad Sci U S A. 2008;105:11076-80. doi: 10.1073/pnas.0804538105.
- [15] Giraudeau B, Leyrat C, Le Gouge A, Léger J, Caille A. Peer review of grant applications: a simple method to identify proposals with discordant reviews. PLoS One. 2011;6:e27557. doi: 10.1371/journal.pone.0027557.
- [16] Graves N, Barnett AG, Clarke P. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel.

- BMJ. 2011;343:d4797. doi: 10.1136/bmj.d4797.
- [17] Abdoul H, Perrey C, Amiel P, Tubach F, Gottot S, Durand-Zaleski I, et al. Peer review of grant applications: criteria used and qualitative study of reviewer practices. *PLoS One*. 2012;7:e46054. doi: 10.1371/journal.pone.0046054.
- [18] Mutz R, Bornmann L, Daniel HD. Heterogeneity of inter-rater reliabilities of grant peer reviews and its determinants: a general estimating equations approach. *PLoS One*. 2012;7:e48509. doi: 10.1371/journal.pone.0048509.
- [19] Schroter S, Groves T, Højgaard L. Surveys of current status in biomedical science grant review: funding organisations' and grant reviewers' perspectives. *BMC Med*. 2010;8:62. doi: 10.1186/1741-7015-8-62.
- [20] Fuhrer MJ, Grabojs M. Grant application and review procedures of the National Institute of Handicapped Research: survey of applicant and peer reviewer opinions. *Arch Phys Med Rehabil*. 1985;66:318-21.
- [21] 白楽ロックビル. ロックビルのバイオ政治学講座 アメリカの研究費とNIH. 東京：共立出版；1996.
- [22] 菅裕明. 切磋琢磨するアメリカの科学者たち—米国アカデミアと競争的資金の申請・審査の全貌. 東京：共立出版；2004.
- [23] Hodgson C. Evaluation of cardiovascular grant-in-aid applications by peer review: influence of internal and external reviewers and committees. *Can J Cardiol*. 1995;11:864-8.
- [24] Hodgson C. How reliable is peer review? An examination of operating grant proposals simultaneously submitted to two similar peer review systems. *J Clin Epidemiol*. 1997;50:1189-95.
- [25] Mayo NE, Brophy J, Goldberg MS, Klein MB, Miller S, Platt RW, et al. Peering at peer review revealed high degree of chance associated with funding of grant applications. *J Clin Epidemiol*. 2006;59:842-8.
- [26] Fogelholm M, Leppinen S, Auvinen A, Raitanen J, Nuutinen A, Väänänen K. Panel discussion does not improve reliability of peer review for medical research grant proposals. *J Clin Epidemiol*. 2012;65:47-52. doi: 10.1016/j.jclinepi.2011.05.001.
- [27] Williams GH, Wara DW, Carbone P. Funding for patient-oriented research. Critical strain on a fundamental linchpin. *JAMA*. 1997;278:227-31.
- [28] Kotchen TA, Lindquist T, Malik K, Ehrenfeld E. NIH peer review of grant applications for clinical research. *JAMA*. 2004;291:836-43.
- [29] Kotchen TA, Lindquist T, Miller Sostek A, Hoffmann R, Malik K, Stanfield B. Outcomes of National Institutes of Health peer review of clinical grant applications. *J Investig Med*. 2006;54:13-9.
- [30] Martin MR, Lindquist T, Kotchen TA. Why are peer review outcomes less favorable for clinical science than for basic science grant applications? *Am J Med*. 2008;121:637-41. doi: 10.1016/j.amjmed.2008.03.031.