

- ・研究課題は、その科学的環境に固有の利点、被験者数、協力関係を享受できるか？

②補足的審査基準

上述の5つの審査基準とは別に「補足的審査基準」として以下の基準が設定されている。これらについては採点はしないが、コメントが記載され、全体の評価の際に考慮される。

○被験者の保護（必須）

- ・被験者が存在するか？
- ・（該当する場合）「被験者のリスク」、「リスクからの保護策の妥当性」、「被験者等に対する潜在的な便益」、「研究から得られる知見の重要性」、「臨床試験におけるデータと安全性に関するモニタリング」に関して評価する。全ての基準が満たされる場合、「リスクは容認できる、または保護が妥当である」と記入する。また1つ以上の基準が適正でない場合、「リスクを容認できない」と記入し、その理由を記述する。

○女性・少数民族・未成年の研究への参加（被験者が存在する場合に必須）

- ・「両性、女性のみ、男性のみ、性別構成不明」のいずれかを記入し、それが容認できるかどうかをコメントする。
- ・「少数民族と多数民族、少数民族のみ、多数民族のみ、民族構成不明、外国人のみ」のいずれかを記入し、それが容認できるかどうかをコメントする。
- ・「未成年と成人、未成年のみ、未成年不参加、未成年参加不明」のいずれかを記入し、それが容認できるかどうかをコメントする。

○脊椎動物の管理と使用

- ・脊椎動物を使用するか？
- ・（該当する場合）「動物実験で使用する種、系統、年齢、性別、頭数」、「動物を使用することの正当性、使用する種と頭数の適

切性」、「獣医療・ケアの適切性」、「鎮痛剤、麻酔、精神安定剤、固定具の使用を含めて、不快感、苦痛、負傷を正常な研究のためにやむを得ない水準にまで抑える手順」、「安楽死の手順、及びアメリカ獣医師会の安楽死指針に適合しない場合、その手順を選択した理由」についてコメントする。

○バイオハザード（必須）

- ・研究で扱う物質や手順が研究者や環境に有害になりうるか？
- ・（該当する場合）適切な保護策が提示されているかどうかをコメントする。

○申請書の再提出

- ・申請書は再提出されたものか？（申請書の再提出は1回まで認められている）
- ・（該当する場合）前回のIRGの評価会議におけるコメントへの回答と変更点を考慮に入れてコメントする。

○申請の更新

- ・申請書は、研究期間が終了して研究の継続を要求するものか？
- ・（該当する場合）前回の研究期間における進捗状況や研究成果を考慮に入れてコメントする。

○申請の修正

- ・申請書は、研究期間の途中で研究計画等の変更（研究期間の延長、研究規模の拡大等）の承認を要求するものか？
- ・（該当する場合）修正の妥当性を考慮に入れてコメントする。

③補足的審査検討項目

さらに「補足的審査検討項目」として、以下の項目が設定されている。これらについて

は採点もせず、全体の評価の際にも考慮されないが、必要に応じてコメントが記述される。

- ・ 予算と研究期間…提案されている研究計画等に照らして正当で合理的かどうか？
- ・ 海外の研究機関からの申請…国内では容易に利用できないような、あるいは国内の既存の資源と補完関係にあるような、海外の優れた技能、資源、人材、環境条件を利用することのメリットが提示されているか？
- ・ 資源共有計画…データの共有、モデル生物の共有、ゲノムの共有を行う計画があるか？ また共有しない場合の理由は正当か？
- ・ Select Agent (指定された有害物質・病原体) …「研究で使用される Select Agent」、 「Select Agent を使用する全ての組織の登録状況」、 「Select Agent の所有、使用、移動を監視するための手順」、 「生物学的安全性、生物学的封じ込め、Select Agent の安全な保管に関する計画」について評価する。

④全体的影響力 (overall impact)

主要審査基準の点数とコメント、補足的審査基準のコメントを参考にして全体的影響力 (overall impact) の評価点数を採点し、コメントを記述する。なお、全体的影響力の評価点数は、各基準の点数の平均値である必要はなく、全ての基準の評価が平均的に優れていなくても、特定の基準の評価が特に優れていれば、全体評価を高くすることが可能である。つまり、申請された研究課題の目的や内容によって重要な基準あるいは重要でない基準があり、査読者の判断による重みづけによって全体評価をすることが可能になっている。

(3) 採点の基準

5つの主要審査基準及び全体的影響力の点数は1~9点で配点され、点数の低い方が評価が高くなるように設定されている。具体的な採点基準は以下のとおりである。

- ・ 1点 特別に優れている (Exceptional) …群を抜いて優れていて、本質的に欠点がみあたらない
- ・ 2点 傑出している (Outstanding) …きわめて優れていて、欠点は無視できる程度である
- ・ 3点 優秀である (Excellent) …とても優れているが、小さな欠点がわずかにある
- ・ 4点 非常によい (Very Good) …優れているが、小さな欠点がいくつかある
- ・ 5点 よい (Good) …優れているが、中程度の欠点が少なくとも1つある
- ・ 6点 十分である (Satisfactory) …ある程度優れているが、中程度の欠点がいくつかある
- ・ 7点 普通である (Fair) …ある程度優れているが、重大な欠点が少なくとも1つある
- ・ 8点 不十分である (Marginal) …少しは優れているが、重大な欠点もいくつかある
- ・ 9点 劣っている (Poor) …優れた点はほとんどなく、重大な欠点が多数ある

なお「小さな欠点 (Minor Weakness)」は、研究のインパクトを実質的に減少させることのない、比較的容易に解決できる欠点、「中程度の欠点 (Moderate Weakness)」は、研究のインパクトが減少してしまうような欠点、「重大な欠点 (Major Weakness)」は、研究のインパクトを重度に限定してしまうような欠点、と定義される。

5. 交付決定後の流れ

研究課題の代表者 (Project

Director(s)/Principal Investigator(s) : PD/PI)は、各種報告書の提出、研究計画の適切な遂行、研究費の管理に関する責任をもつ。実際の研究費の運用管理は所属施設が行うこととなり、PD/PIは経理担当者と十分に相談した上で研究費の適正な執行を行う必要がある。定期的に提出が義務づけられている報告書として、進捗報告(研究課題の進捗状況等)、発明報告(特許等の知的財産権の出願・登録状況等)、財務報告などがある。

研究課題の遂行にあたっては、NIHの各研究所・センターに配置されている Program Official (PO)、Grants Management Officer (GMO)に随時相談することができる。POは研究費のプログラムの側面、つまり科学的・技術的側面の責任をもち、GMOは研究費の実務・管理面の責任をもち、それぞれの立場から研究課題の円滑な遂行に必要な助言を受けることができる。

POの具体的な業務としては、所属する研究所・センターのミッションに適合した研究イニシアティブや研究トレーニングプログラムの開発、CSRやSROとの連携・調整、進捗報告のレビュー、サイトビジットへの参加など、研究費交付後の研究課題に対する進捗管理や支援に関する様々な活動を行っている。

D. 考察

①「二段階」の研究評価システム

NIHの新しい研究評価システムのわが国と異なる特徴として、申請書の審査・評価が「二段階」で実施される点が挙げられる。NIHの研究予算は各研究所・センターによって運営管理され、内部研究 (Intramural Research) と外部研究 (Extramural Research) に配分されるが、外部研究の申請に関して二段階評価が採用されている。つまり、各研究所・センターによる研究開発の方針の設定、その方針

にしたがった研究課題の公募、申請書の受付、申請書の評価、研究課題の採択の一連のプロセスにおいて、研究課題の公募、申請書の受付、一次評価 (Initial Level of Review) を Center for Scientific Review (CSR) が NIH 横断的に実施し、二次評価 (Second Level of Review) と研究課題の採択を各研究所・センターが行う、という仕組みである。一次評価は、CSRが運営管理する、特定の学問領域、専門分野の学識経験者で構成される複数の統合審査グループ (Integrated Review Group : IRG) によって実施され、主に「学術的・科学的観点」から申請書を評価する。二次評価は、各研究所・センターの諮問委員会によって実施され、一次評価の結果を踏まえた上で、主に「行政的・政策的観点」から申請書を評価し、採択課題を決定する。

それに対して、厚生労働科学研究費補助金では一段階の事前評価のみで採択課題が決定されている。事前評価委員会もNIHのIRGも研究事業(研究課題)に関連する学問分野の学識経験者で構成されているが、事前評価委員会には学術的側面だけでなく行政的側面からの評価が求められている。学識経験者は必ずしも行政施策に精通しているわけではなく、研究課題の行政的、政策的な有用性や緊急性を十分に評価することは困難であると考えられる。また学術的側面と行政的側面を同時に評価することによって、両側面を「平均」した形で総合評価されてしまう可能性があり、学術的側面の評価が低い課題でも行政的評価が高ければ採択されてしまう可能性がある。さらに、事前評価委員会は実質的に研究課題の採択の責任を有しているため、行政施策への反映を重視する厚生労働科学研究費補助金に適した研究課題として、学術的側面よりも行政的側面が重視されてしまう可能性がある。したがってわが国においても、一次評価で学術的価値の高い申請課題を「スクリーニング」して、学術的価値を保証した上で、二次

評価で行政的・政策的価値の高い研究課題を採択する、という NIH のシステムを導入することによって、学術的側面と行政的側面のともに評価の高い研究課題を採択することが可能であると考えられる。

②研究事業全体に共通する研究評価システム

上述したように、NIH では、各研究所・センターが外部研究資金 (Extramural Research) を運営管理しているが、それに関して指導、監督、助言を行う Office of Extramural Research (OER)、研究課題の公募、申請書の受付、一次評価の運営管理を行う CSR が「NIH 横断的」に設置されている。

厚生労働科学研究費補助金の「研究事業」は厚生労働省の所管課室によって運営管理される外部研究資金と捉えることができる。各研究事業の運営管理業務の中には、申請書の受付、評価、採択、交付決定、各種報告書の受理など、共通する部分が多く、これらの業務は各研究事業の所管課室で別々に実施するよりも一括して遂行した方が効率的であると考えられる。これらの業務のうち、申請書の受付に関しては府省共通研究開発管理システム (e-Rad) によって、研究報告書概要版の受理に関しては厚生労働科学研究成果データベースによって、すでに研究事業横断的に実施されている。今後は、申請書の評価に関する共通したシステムを構築することによってさらに効率的な事務の遂行が可能になると考えられる。具体的には、CSR の IAR (Internet Assisted Review) のような電子研究評価システムの開発・運用、CSR の科学審査官 (Scientific Review Officer : SRO) のような、評価委員会の運営管理を行う専門職員の配置などが挙げられる。

③申請課題の評価基準と評価の手続き

NIH における申請課題の評価では、「主要審査基準」、「補足的審査基準」、「補足的

審査検討項目」を評価基準として、それぞれについて評価 (主要審査基準では採点とコメントの記述、補足的審査基準ではコメントの記述) を行った上で、それらを参考にして「全体的影響力 (overall impact)」の総合評価 (採点とコメントの記述) を行っている。また主要審査基準及び全体的影響力の点数は 1~9 点で配点され、欠点の多さや重大さなどの具体的な採点基準が設定されている。

厚生労働科学研究費補助金においても「評価にあたり考慮すべき事項」が設定されているが、それぞれに採点やコメントは求められていないため、総合評価の際にそれらの基準が十分に考慮されていない可能性がある。したがってわが国においても、評価委員の負担が大きくなる範囲で、各事項で採点とコメントの記述を行った上で総合評価を行う方がよいと考えられる。また点数の配点に関しても、全体的に高く評価するなど、評価委員によって採点の傾向が異なる場合が多い。したがってわが国においても明確な採点基準を設定して、評価者間の採点の傾向のばらつきを縮小させる必要がある。

ただし総合評価にあたっては、NIH では、全ての基準の評価が平均的に優れていなくても、特定の基準の評価が特に優れていれば総合評価を高くすることが可能になっている。これは、公募課題の目的や内容によって評価基準の優先順位や重要度が異なることに対応するためであると考えられる。したがってわが国で適用する際には、評価者に対して公募課題の採択方針等を十分に理解してもらった上で、評価者の判断による重みづけで総合評価をしてもらう必要がある。

NIH では、主要審査基準として「重要性 (Significance)」(研究テーマの医学・社会における重要性、研究成果が医学・社会に及ぼすインパクト等)、「研究者 (Investigator(s))」(研究者及び研究班の研究遂行能力等)、「革新性 (Innovation)」

(研究テーマ・研究成果の新規性等)、「アプローチ (Approach)」（研究計画の適切性・妥当性、研究によるリスクの管理等）、「環境 (Environment)」（研究機関の資源や支援体制、研究者・研究課題が利用できる研究環境等）が設定されている。わが国の研究計画書でもほとんど同様の内容の記載が求められているが、「環境」に関する記載が十分でないのが現状である。NIHの申請書では、研究者の所属研究機関において研究課題で利用可能な施設（研究施設、動物実験施設等）や設備（実験機器、測定機器等）の状況を記載することが求められており、それが評価にも反映されている。研究成果を産出するためのインプット（生産要素）には、施設、設備などの固定的生産要素と、消耗品、スタッフなどの可変的生产要素がある。研究課題を遂行するための「可変的」な研究資金が十分にあったとしても、それを遂行するための環境、特に研究機関の施設・設備、支援体制などの「固定的」な資源が整備されていなければ、研究費の効果的かつ効率的な運用は困難であると考えられる。したがってわが国においても、環境、特に研究機関の施設・設備、及び申請課題に対する支援体制を評価の要素として重視する必要がある。

④中間・事後評価

NIHでは、中間評価、事後評価といった、研究期間の途中や終了時に評価する枠組みが設定されていない。研究期間の途中においては、研究計画等の変更に関する申請があった場合、「申請の修正」としてこれまでの進捗状況が評価される。また研究期間の終了時には、さらなる研究の継続に関する申請があった場合、「申請の更新」として前回の研究期間における研究成果が評価される。しかしこれらは全ての研究課題で実施されるわけではない。一方、厚生労働科学研究費補助金では、NIHや科学研究費補助金とは異なり、

事前評価、中間評価、事後評価と、研究期間の途中で複数回の評価を実施する仕組みとなっている。

研究期間内に複数回の評価を実施することによって、研究の進捗状況が適時に評価され、研究期間の途中で研究の質を向上させることが可能である反面、評価のための資料作成や申請書の査読など、研究者、評価委員の双方の負担が大きくなるという問題がある。一方、研究期間の途中で評価がない場合、進捗状況や研究成果に関わらず研究期間内の研究費は確保されるため、研究の質を向上させるインセンティブは小さくなる可能性がある。しかしほとんどの研究者は研究期間が終了した後も研究を継続したいと考えており、そのためには当該研究課題で高い研究成果を産出した上で申請の更新を行う必要がある。したがって研究期間の途中で評価がなかったとしても、多くの研究者、研究課題は当初の研究計画にしたがって高い研究成果を得ることができると考えられる。そして中間評価の効果は、研究に対する意欲の低い研究者、質の低い研究課題に対して、厳しい評価を行うことによって若干の改善を促す程度の限定的なものにとどまると考えられる。

しかし、優れた研究計画であっても、研究期間の途中で様々な障害に直面し、研究を円滑に遂行できない状況に陥る可能性がある。そのような場合に必要となるのは「研究評価」ではなくむしろ「研究支援」であり、NIHではそのための専門職員として Program Official (PO) を配置している。POは各研究所・センターの職員で、研究事業の科学的・技術的側面の責任をもち、新しい研究イニシアティブの開発、研究評価に関するCSRやSROとの連携・調整、各研究課題の進捗報告のレビューやサイトビジットなど、研究費交付後の研究課題の進捗管理や支援に関する様々な活動を行っている。

つまり、NIHでは、申請課題に対する厳密な「研究評価」によって質の高い研究課題を採択し、研究期間内におけるP0等による充実した「研究支援」によって研究課題の円滑な遂行を促進することによって、最終的に価値の高い研究成果を産出することを目指していると考えられる。それに対して厚生労働科学研究費補助金では、これらの全ての機能と役割を事前評価、中間評価、事後評価の「研究評価」が担っていると考えられる。どちらのシステムも一長一短であるが、NIHのように研究期間の途中での評価を行わないシステムを採用する場合は、P0の量的確保と機能の充実が必要条件となると考えられる。

⑤研究評価の実施時期

NIHでは、研究課題の公募要項（Funding Opportunity Announcement：FOA）の公布と申請の受付を「随時」行うとともに、申請書の受理、一次評価、二次評価、交付決定の一連の研究評価プロセスに少なくとも約半年間の期間を要している。それに対して厚生労働科学研究費補助金は、10月末の公募要項の公布、12月初めまでの申請書の受理、3月末の交付決定と、特定の時期に集中して、比較的短期間に研究評価を遂行する必要がある。そのため、特定の時期に評価委員や事務局の負担が大きくなるという問題がある。

わが国とアメリカでは政府の会計システムが異なるため研究課題を年度途中から開始することはできないが、研究評価の時期を変更することは可能であると考えられる。厚生労働科学研究費補助金では、研究開始時の「事前評価」、研究途中の「中間評価」、研究終了時の「事後評価」が年度末のほとんど同じ時期に実施されているが、そのうち中間評価、事後評価に関しては時期を移動させることが可能であると考えられる。

中間評価に関しては、研究開始2年目の中ごろに実施し、その評価結果を踏まえて3年

目の継続の可否を検討することが考えられる。この場合、2年目の継続の可否に関する評価はできないが、1年間で研究成果を上げるとは困難な場合が多く、研究成果に関する評価が十分にできないのが現状である。したがって実質的に2年間はほとんど無条件で継続することを認める一方で、1年半後の中間評価で十分な研究成果が得られていない相当数の課題については3年目の継続を認めないという形で評価することによって、真に価値の高い課題のみに研究費を継続して配分することが可能になると考えられる。

また事後評価に関しては、研究開始3年目の中ごろに実施し、研究期間全体の評価を行うことが考えられる。研究期間は半年間残されているが、この時期はこれまでの研究成果をとりまとめる時期であり、新しい研究成果が得られることは少ないと考えられる。そのため研究開始の2年半後の評価で研究期間全体の評価がほぼ可能であると考えられる。また当該研究課題がさらなる継続を求めて新規に申請した際には、年度の間で確定した事後評価の結果を、年度末に実施される事前評価に活用することが可能であると考えられる。

⑥公募課題の範囲

NIHの公募要項には、特定の研究テーマに限定せずに研究者の自由な自発的な発想に基づく研究計画を募集する「Parent Announcement」、特定の研究領域において幅広く研究計画を募集する「Program Announcement（PA）」、特定の研究テーマに対して特定の研究費の範囲内で実施する研究計画を募集する「Request for Application（RFA）」の3種類がある。

わが国と比較すると、科学研究費補助金はParent Announcementに、厚生労働科学研究費補助金はRFAにそれぞれ近いと考えられる。RFAは、設定された研究の目的や目標を確実に達成することが可能である反面、研究者の

裁量の余地が少なく、新しいアイデアや知見が得られにくいという問題もある。一方 Parent Announcement は、研究者の関心のみで申請されるため、行政的・政策的ニーズに適合した研究課題を採択することが困難である。したがってわが国でも、PA のように、研究の領域や行政的・政策的な課題を緩やかに設定した上で、研究者の自由な発想に基づく研究計画を申請できる枠組みを設定することによって、行政ニーズに適合した研究テーマにおける新しい知見やイノベーションの創出を促す必要がある。

⑦研究評価、研究支援に関わる専門職員

NIH では、研究評価、研究支援を実施するに当たって、複数の専門職員が配置されている。具体的には、IRG のレビュー・評価のプロセスの運営管理の責任をもつ科学審査官 (Scientific Review Officer : SRO) 、研究プロジェクトの科学的・技術的側面の責任をもつ Program Official (PO) 、研究費の実務・管理面の責任をもつ Grants Management Officer (GMO) などが挙げられる。

国立保健医療科学院等の厚生労働科学研究費補助金の FA では、PO と交付事務担当者が配置されているが、いずれも専門職員として位置づけられておらず、PO に関しては研究職との兼務である。また国立保健医療科学院の PO は、研究動向の把握、研究課題評価の運営管理、研究課題への支援、研究成果のとりまとめなどの業務を実施しているが、これらは NIH における SRO と PO の業務に相当する。NIH で複数の専門職員が実施する業務を研究職の兼務である PO が実施するのは不可能である。したがって、厚生労働科学研究費補助金の研究評価、研究支援の機能を強化するためには、少なくとも PO を専門職員とすること、可能であれば研究評価の専門職員と研究支援の専門職員をそれぞれ別に配置することが必要である。

⑧研究費配分機能 (FA) のあり方

NIH では、各研究所・センターのミッションや研究開発戦略を遂行するために研究資金が配分されているが、それを内部研究 (Intramural Research) 、外部研究 (Extramural Research) のどちらに、どの程度配分するかは、各研究所・センターの目的や目標を達成するためにどちらで実施する方が効率的か、という観点で決定される。それに対して、厚生労働科学研究費補助金の FA は外部研究資金の配分機能のみが移管されており、内部研究資金は各機関の運営費として別の予算から配分されている。しかし研究機関の専門領域に関係する研究事業の FA 機能が移管されているため、FA が配分する外部研究資金をその研究機関の研究者が獲得してしまう場合がある。その結果、外部に配分すべき研究資金を内部に環流するという不自然な流れが発生してしまい、外部研究資金の本来の役割を果たしていないという問題がある。

したがってわが国においても、NIH の研究所・センターと同様に、FA の研究機関に研究事業の予算を内部研究と外部研究に配分する権限を付与する必要がある。それによって、FA の研究機関自らが実施すべき研究内容 (例えば、行政的ニーズが重視される研究テーマなど) と大学等の外部の研究機関が実施する方が効率的である研究内容 (学術的成果が期待される研究テーマなど) を明確に設定することができ、FA の研究機関と大学等との役割分担や連携が円滑に推進されることが考えられる。ただしそのためには、NIH の研究所・センターのように、FA の研究機関のミッションや研究開発戦略を明確に設定する必要がある。

⑨厚生労働科学研究費補助金の使途の範囲

研究課題に配分される研究費は本来、消耗品、スタッフなどの可変的生産要素に投入される資金であり、施設・設備などの固定的生

産要素は研究機関がそれとは別に整備すべきである。しかし現状では、実験機器、測定機器などの設備に関して、正当な理由があれば研究費で購入できるようになっている。その正当な理由は「当該研究課題のみに使用すること」であるが、実際には他の研究課題でも使用可能であり、また施設・設備を中心に考えれば、特定の研究課題のみに使用することは非効率的である。

したがって、研究費の適正な執行の観点から、研究費で施設・設備を購入できないようにすること、その代わりに研究費とは別の枠組みで施設・設備を充当する仕組みを構築することが必要である。現在のところ間接経費がその役割を果たすと考えられるが、厚生労働科学研究費補助金では、科学研究費補助金とは異なり、間接経費としての予算が確保されていないこと、間接経費の使途に関する明確な規定がないことなど、施設・設備を充実する機能を十分に果たしていない可能性がある。

E. 結論

効果的かつ効率的な研究評価システムの具体的な構造と機能を検討することを目的として、アメリカの National Institutes of Health (NIH) における研究評価システムの実態に関して、文献データベース、Web 等を用いて、関連する文献、報告書、資料などを網羅的に収集し、その運営管理体制、研究課題の評価基準等に関する情報を抽出し、わが国への適用可能性を検討した。その結果、以下のことが明らかとなった。

①NIH の外部研究 (Extramural Research) における申請課題の評価は「二段階」で実施されている。一次評価は Center for Scientific Review (CSR) において実施され、主に「学術的・科学的観点」から評価される。二次評価は NIH の各研究所・セン

ターにおいて実施され、一次評価の結果を踏まえた上で、主に「行政的・政策的観点」から評価され、採択課題が決定される。わが国においても、一次評価で学術的価値の高い申請課題を「スクリーニング」して、二次評価で行政的価値の高い研究課題を採択する、という「二段階評価システム」を導入することによって、学術的側面と行政的側面の両面の評価の高い研究課題を採択することが可能になることが示唆された。

②一次評価の評価基準として、主要審査基準、補足的審査基準、補足的審査検討項目が設定され、それぞれの基準について評価を行った上で、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。評価点数は 1~9 点で配点され、欠点の多さや重大さなどの具体的な採点基準が設定されている。また「主要審査基準」として重要性 (Significance)、研究者 (Investigator(s))、革新性 (Innovation)、アプローチ (Approach)、環境 (Environment) が設定されている。わが国においても、申請課題の評価にあたっては、研究計画、研究者だけでなく、それを取り巻く「研究環境」、具体的には研究機関の施設・設備、支援体制などの「固定的」な資源の状況を把握・評価する必要があることが示唆された。

③NIH 横断的な組織として、各研究所・センターの外部研究資金の運営管理に関して指導、監督、助言を行う Office of Extramural Research、公募要項 (Funding Opportunity Announcement) の公布、申請書の受付、一次評価の運営管理を行う CSR が設置されている。また研究評価、研究支援に関わる専門職員として、一次評価の運営管理 (査読者の選定、評価委員会の開催、評価結果のとりまとめなど) の責任をもつ科学審査官

(Scientific Review Officer : SR0) 、研究プロジェクトの科学的・技術的側面の責任をもち、新しい研究イニシアティブの開発、研究課題の進捗管理などの研究支援を行う Program Officer (PO) 、研究費の実務・管理面の責任をもつ Grants Management Officer (GMO) などが配置されている。わが国においても、研究評価・研究支援の機能の強化と効率化を図るために、厚生労働科学研究費補助金全体に共通する申請書の評価業務を支援する電子研究評価システムの開発・運用、Program Officer (PO) の専門職員としての位置づけの明確化、研究評価と研究支援の専門職員の独立した配置などが必要であることが示唆された。

④NIH では、研究期間の途中での中間評価、研究期間終了時の事後評価は明確には実施されていないが、各研究所・センターの PO 等による研究期間を通じた充実した「研究支援」によって研究課題の円滑な遂行を促進している。この点を踏まえて、わが国の厚生労働科学研究費補助金の研究評価、特に中間・事後評価に関して、より質の高い研究課題を継続させるため、また研究者、評価委員、事務局の評価に係る負担を軽減させるために、以下のような研究評価サイクルが提案された。

- ・申請者は前年度の 12 月に研究計画書（新規用）を提出する。
- ・事前評価委員会は前年度の 2～3 月に「事前評価」を実施し、採択の可否を審査する。
- ・研究者は、研究開始 1 年目の 12 月に研究計画書（2 年目継続用）を提出する。
- ・PO は研究計画書（研究の進捗状況、次年度の研究計画等）をレビューし、研究の遂行上特に重大な問題がない研究課題については 2 年目の継続を承認し、問題のある課題については 2～3 月に開催される事前評価

委員会に諮問する。事前評価委員会は前年度の研究計画（新規用）の採択の決定に関する妥当性を検証するとともに、継続の可否を審査する。

- ・研究者は研究開始 2 年目の 8 月に中間結果報告書（中間的な研究成果、次年度の研究計画等）を提出する。
- ・中間・事後評価委員会は研究開始 2 年目の 10～11 月に「中間評価」を実施し、継続の可否を審査する。この際に、十分な研究成果が得られている、あるいは得られることが期待される課題のみを継続として、それ以外の相当数の課題は 2 年で終了とする。
- ・継続が承認された研究課題の研究者は研究開始 2 年目の 12 月に研究計画書（3 年目継続用）を提出する。
- ・PO は中間評価のコメントにしたがって研究計画等が修正、改善されているかどうかを確認し、3 年目の継続を承認する。修正、改善が十分でない課題に対しては再提出を求める。
- ・研究者は研究開始 3 年目の 8 月に最終結果報告書（最終的な研究成果、次年度の新規申請の予定等）を提出する。
- ・中間・事後評価委員会は研究開始 3 年目の 10～11 月に「事後評価」を実施し、研究成果の評価を行う。
- ・PO は事後評価のコメントにしたがって事後評価報告書を作成し、当該研究課題が次年度に新規申請を行った場合に 2～3 月に実施される事前評価委員会に提出する。
- ・事前評価委員会は、提出された最終結果報告書と次年度に向けて提出された研究計画書（新規用）を資料として「事前評価」を実施し、採択の可否を審査する。

F. 健康危険情報 なし

G. 研究発表 なし

H. 知的財産権の出願・登録状況 なし

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）
分担研究報告書

健康安全・危機管理対策に関連する研究評価・研究支援方法の開発
—研究課題の評価点数の状況とその関連要因の検討—

研究分担者 武村 真治（国立保健医療科学院公衆衛生政策部地域保健システム室長）

研究要旨

目的：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）の研究課題の評価（事前評価、中間評価、事後評価）の傾向とその影響要因を明らかにするとともに、研究成果との関連から研究評価の妥当性を検証することによって、研究評価が研究成果に結びつくような、推進すべき研究課題に対して遂行能力の高い研究者・研究機関が適切に評価されるような研究評価手法を開発する。

方法：対象は平成18年度～21年度に本研究事業の交付を受け、研究を終了した研究課題61課題であった。使用したデータは、事前評価、中間評価（研究開始1年目、2年目）、事後評価の評点（専門委員による専門点、行政委員による行政点、全体の評価点）、研究費総額、研究課題の進捗状況（研究事業推進官（Program Officer：PO）への研究班会議の開催通知の有無、研究班会議へのPOの出席の有無）、研究成果（原著論文及びその他論文（和文、英文等）の発表件数、国内・国際学会での発表件数、施策への反映、普及・啓発活動等）などであった。事前評価、中間評価、事後評価の評点の変化と関連、研究課題の進捗状況と評点との関連、評点及び研究費交付額と研究成果との関連等を分析した。

結果：専門点は事前評価と比較して事後評価が低く、行政点と評価点は事前評価と比較して中間評価、事後評価が低かった。事前評価、中間評価、事後評価の間の相関は、専門点では事前と事後、中間（1年目）と中間（2年目）、中間（1年目）と事後、中間（2年目）と事後の間で正の相関がみられた。行政点では事前と中間（1年目）、中間（1年目）と事後、中間（2年目）と事後の間で正の相関がみられた。評価点では事前と中間（1年目）、事前と事後、中間（1年目）と中間（2年目）、中間（1年目）と事後、中間（2年目）と事後の間で正の相関がみられた。専門点と行政点の間の相関は、事前評価、中間評価においてはみられなかったが、事後評価において正の相関がみられた。研究分野別では「地域健康危機管理の基盤形成に関する研究分野」で、事前評価の専門点、行政点、評価点、事後評価の行政点と評価点が低かった。

研究班会議の開催通知があった研究課題の割合は、研究開始1年目で約8割、2年目で約6割、3年目で約7割、POが出席した研究課題の割合は、1年目で約6割、2年目で約4割、3年目で約6割と、いずれも2年目に減少し、3年目に増加していた。またPOが出席した研究課題は事後評価の行政点、評価点が高かった。

「地域健康危機管理の基盤形成に関する研究分野」、「健康危機管理・テロリズム対策システム研究分野」は英文等による論文の発表件数、国内・国際学会での発表件数といった学術的成果が少なかった。また特許の出願・取得がなかった課題が9割以上、施策への反映がなかった課題が約7割、普及・啓発活動がなかった課題が約5割と、行政的成果が非常に少なかった。学術的成果の変数の間、行政的成果の変数の間では互いに正の相関がみられたが、学術的成果と行政的成果の間の相関はほとんどみられなかった。

研究費総額は学術的成果との正の相関がみられたが、行政的成果との相関はみられなかった。

研究成果と評点との関連では、事前評価は国内学会での発表以外の学術的成果との相関がみられなかった。中間評価（1年目）は全ての学術的成果との相関がみられなかった。中間評価（2年目）は、原著論文（和文、英文等）との相関はみられなかったが、その他論文（和文、英文等）、国内・国際学会での発表との正の相関がみられた。事後評価はその他論文（和文）以外の学術的成果との正の相関がみられた。また行政的成果に関しては、施策に反映した研究課題、普及・啓発活動を実施した研究課題の方が事前評価が低かった。中間評価はほとんどの行政的成果との相関がみられなかった。事後評価は全ての行政的成果との相関がみられなかった。

結論：研究評価の結果に基づいた研究支援方策として、①全ての研究課題に対して、事前評価の段階で期待されていた学術的・行政的成果を十分に認識した上で研究を遂行できるように支援すること、②事前評価や中間評価の低い研究課題に対して、特に学術的側面から、研究期間全体を通じて優れた成果を得られるように支援すること、③行政点の低い研究課題に対して、研究期間の途中段階であっても、行政ニーズに応じた適時の成果を比較的短期間に得られるように支援すること、④「地域健康危機管理の基盤形成に関する研究分野」の研究課題、研究開始2年目の研究課題、POが班会議に出席できなかった研究課題に対して重点的に支援すること、が必要であることが示唆された。

研究評価の研究成果に対する基準妥当性に関しては、同時妥当性（同時に測定される基準との相関）は学術的成果に対しては高いが行政的成果に対しては低いこと、予測妥当性（将来測定される基準との相関）は学術的成果、行政的成果ともに低いことが示された。したがって基準妥当性を改善するためには、①研究報告書や研究計画書の様式に行政的成果を記載する項目を設定し、それらを具体的かつ詳細に記載するように研究代表者に周知すること、②研究計画書の「期待される成果」を具体的に記述するように申請者に周知すること、③ヒヤリング等を通じて、研究計画書の書面には顕在化しにくい研究に対する意欲や積極的な姿勢などの「意志の力」を測定・評価すること、が必要であることが示唆された。

A. 研究目的

研究資金を効率的に運用する上で、「事前評価」において研究計画が適正に評価され、「中間評価」において研究の進捗状況が適正に評価され、「事後評価」において研究成果が適正に評価されるような、研究評価システムを確立することが重要である。このシステムが円滑に運営されることによって、より多くの研究成果が期待される研究課題が採択され、その課題の調査研究が円滑に推進され、最終的に期待された（以上の）研究成果が産出されることになる。

平成18年度より、国立保健医療科学院は、厚生労働科学研究費補助金の研究事業に係る研究費配分機能（Funding Agency：FA）を担

っており、評価委員の選定、評価委員会の運営、評価委員会への説明と報告、評価結果やコメントの取りまとめなどの業務を通じて、研究評価システムの円滑な遂行を支援している。しかし現状の研究評価システムが適正に運営されているかどうか明らかではなく、これを検証するためには、現在の研究評価の実態と傾向を分析し、その問題点を明らかにする必要がある。

そこで本研究は、健康安全・危機管理対策総合研究事業の交付を受けた研究課題の評価（事前評価、中間評価、事後評価）の傾向とその影響要因を明らかにするとともに、研究成果との関連から研究評価の妥当性を検証することによって、研究評価が研究成果に結び

つくような、推進すべき研究課題に対して遂行能力の高い研究者・研究機関が適切に評価されるような研究評価手法を開発することを目的とした。

B. 研究方法

対象は、平成18年度～21年度に「健康安全・危機管理対策総合研究事業（地域健康危機管理研究事業、健康危機管理・テロリズム対策システム研究事業）」の交付を受け、研究を終了した研究課題とした。

分析に使用したデータは、研究分野、研究年度、開始年度、終了年度、各年度の評点、研究費総額、各年度の研究課題の進捗状況、研究成果等とした。

研究分野のカテゴリーは、「地域健康危機管理の基盤形成に関する研究分野（以下、「地域」とする）」、「水安全対策研究分野（以下、「水道」とする）」、「生活環境安全対策研究分野（以下、「生活衛生」とする）」、「健康危機管理・テロリズム対策システム研究分野（以下、「テロ」とする）」とした。

評点に関しては、事前評価、中間評価（研究開始1年目）、中間評価（研究開始2年目）、事後評価のそれぞれについて、専門委員による評価点数の平均値（以下、専門点）、行政委員による評価点数の平均値（以下、行政点）、専門点と行政点の平均値（以下、評価点）を使用した。なお、研究期間が1年間の研究課題については、研究開始1年目の評価を事後評価として、中間評価（研究開始1年目）、中間評価（研究開始2年目）を欠損値とした。また研究期間が2年間の研究課題については、研究開始2年目の評価を事後評価として、中間評価（研究開始2年目）を欠損値とした。

平成17年度以前に開始された研究課題に関しては、事前評価、中間評価のデータが得られなかったため、それらの値を欠損値として分析から除外した。

平成20年度以降に開始された研究課題の事前評価に関しては、行政委員による評価が実施されていないため、行政点を欠損値とした。そのため、専門点と評価点は同じ値とした。

研究課題の進捗状況に関しては、平成19年度より、研究班会議の開催を研究事業推進官（Program Officer:P0）に通知してもらい、P0は可能な限り出席して、研究遂行に必要な情報を提供するという活動を実施している。そこで、各研究年次（研究開始後の年数）のP0への研究班会議の開催通知の有無、研究班会議へのP0の出席の有無を進捗状況の変数とした。

研究成果に関しては、「厚生労働科学研究成果データベース」の行政効果報告のデータを用いた。具体的には、原著論文（和文）の発表件数、原著論文（英文等）の発表件数、その他論文（和文）の発表件数、その他論文（英文等）の発表件数、国内学会での発表件数、国際学会等での発表件数、特許の出願及び取得状況（件数）、施策への反映（件数）、普及・啓発活動（件数）のデータを用いた。

分析に関しては、開始年度別、研究分野別に評点及び進捗状況を分析した。次に、事前評価、中間評価（1年目）、中間評価（2年目）、事後評価の評点の変化と相関、研究課題の進捗状況と評点との関連、評点及び研究費交付額と研究成果との関連等を分析した。

（倫理面への配慮）

評点については、公平性・中立性・透明性の確保、税金を財源とする厚生労働科学研究費補助金の適正な使用の観点から、個人情報秘密保持の観点、及び研究者間の新たな利害関係を生じさせないように個人の特長ができないように配慮しつつ、国民に対して評価結果とその反映状況について積極的な情報の提供を図ることが求められている。また、研究者や評価委員の個人情報は取り扱わない。

以上の点から、本研究の倫理的な問題は小さいと考えられた。しかし分析結果の提示にあたっては、個別の研究課題を同定できないように配慮した。なお研究の実施にあたっては、国立保健医療科学院研究倫理審査委員会の承認を得た（承認番号 NIPH-IBRA#09010）。

C. 研究結果

1. 研究課題の概要

対象は、平成 18 年度～21 年度に「健康安全・危機管理対策総合研究事業（地域健康危機管理研究事業、健康危機管理・テロリズム対策システム研究事業）」の交付を受け、研究を終了した研究課題数は 61 課題であった。研究分野別では、地域が 33 課題（54.1%）、水道が 9 課題（14.8%）、生活衛生が 13 課題（21.3%）、テロが 6 課題（9.8%）であった。

表 1 に、研究分野別にみた開始年度を示した。平成 19 年度に開始された研究課題の割合が 39.3%と大きい。これは、平成 19 年度よりテロの研究分野が新設されたためである。

表 2 に、研究分野別にみた研究期間を示した。ほとんどが 3 年間であったが、1～2 年間の短期間で実施される研究課題は 21.3%であった。「水道」は、他の分野と比較して、研究期間が 3 年間である割合が大きかった。

表 3 に、研究分野別にみた、各研究年次の班会議の開催通知があった研究課題、及び P0 が出席した研究課題の数と割合を示した。全体では、開催通知の割合は、研究開始 1 年目で 84.6%、2 年目で 56.8%、3 年目で 70.4%であった。また P0 出席の割合は、1 年目で 57.7%、2 年目で 37.8%、3 年目で 59.3%であった。開催通知、P0 出席いずれも、2 年目に減少し、3 年目に増加していた。

研究分野別では、1 年目に関しては「生活衛生」の開催通知の割合が小さく、「生活衛生」、「テロ」の P0 出席の割合が小さかった。2 年目に関しては、「生活衛生」の開催通知の割合が小さく、「水道」の P0 出席の割合が

大きかった。3 年目に関しては、「地域」の開催通知の割合が小さく、「生活衛生」の P0 出席の割合が大きかった。年次の変化をみると、「地域」では、年次を追うごとに開催通知の割合が減少し、「水道」では年次を追うごとに P0 出席の割合が減少し、「生活衛生」では 3 年目に開催通知の割合が増加し、年次を追うごとに P0 出席の割合が増加し、「テロ」では 3 年目に開催通知、P0 出席の割合が増加した。

表 4 に、研究分野別にみた、研究費総額を示した。全体では、平均値は 3,751.6 万円、中央値は 2,400 万円で、高い方にすそをひく分布であった。1,500 万円未満が約 3 割、2,500 万円未満が約 5 割、5,000 万円未満が約 8 割であった。

研究分野別では、ノンパラメトリック検定の結果、「生活衛生」は他の分野と比較して研究費総額が多かった。なお平均値では「水道」が他の分野と比較して高かったが、2,000 万円以上 2,500 万円未満と 1 億円以上の二峰性の分布であることが影響していると考えられる。

2. 事前評価、中間評価、事後評価の評点の状況

表 5 に、研究分野別にみた、事前評価の評点の平均値と標準偏差を示した。全体の平均値は、専門点 3.69、行政点 3.99、評価点 3.84 であった。専門点に関しては「地域」、「テロ」が低く、「水道」が高かった。行政点に関しては「地域」が低く、「テロ」が高かった。評価点に関しては「地域」が低かった。

表 6 に、研究分野別にみた、中間評価（研究開始 1 年目）の評点の平均値と標準偏差を示した。全体の平均値は、専門点 3.61、行政点 3.72、評価点 3.67 であった。分野間の評点の差はみられなかった。

表 7 に、研究分野別にみた、中間評価（研究開始 2 年目）の評点の平均値と標準偏差を

示した。全体の平均値は、専門点 3.62、行政点 3.69、評価点 3.66 であった。分野間の評点の差はみられなかった。

表 8 に、研究分野別にみた、事後評価の評点の平均値と標準偏差を示した。全体の平均値は、専門点 3.46、行政点 3.74、評価点 3.60 であった。行政点に関しては「地域」が低かった。評価点に関しては「地域」、「テロ」が低かった。

表 9 に開始年度別の事前評価、表 10 に開始年度別の中間評価（1 年目）、表 11 に開始年度別の中間評価（2 年目）、表 12 に開始年度別の事後評価の評点の平均値と標準偏差を示した。いずれに関しても、開始年度で評点の差はみられなかった。

表 13 に、全ての研究課題について、事前評価、中間評価（1 年目）、中間評価（2 年目）、事後評価の、専門点、行政点、評価点の推移を示した。事前評価、中間評価（1 年目）、中間評価（2 年目）、事後評価の組み合わせで対応のある t 検定を行った結果、専門点に関しては、事前評価と比較して事後評価が低かった。また行政点、評価点に関しては、事前評価と比較して、中間評価（1 年目）、中間評価（2 年目）、事後評価が低かった。

表 14 に、研究開始 1 年目の研究課題の進捗状況の別にみた事前評価及び中間評価（1 年目）の評点を示した。班会議の開催通知の有無及び P0 の出席の有無を用いて、「班会議の開催通知がなかった」、「班会議の開催通知があったが、P0 が出席できなかった」、「班会議の開催通知があり、P0 が出席した」の 3 群に分類し、評点を比較した。開催通知や P0 の出席の「前」の事前評価の評点に関しては、群間で差がみられなかった。また開催通知や P0 の出席の「後」の中間評価（1 年目）の評点に関しても、群間で差がみられなかった。

表 15 に、研究開始 2 年目の研究課題の進捗状況の別にみた中間評価（1 年目）と中間評価（2 年目）の評点を示した。開催通知や P0

の出席の「前」の中間評価（1 年目）の評点に関しては群間で差がみられなかった。また開催通知や P0 の出席の「後」の中間評価（2 年目）の評点に関しても群間で差がみられなかった。

表 16 に、研究開始 3 年目の研究課題の進捗状況の別にみた中間評価（2 年目）と事後評価の評点を示した。開催通知や P0 の出席の「前」の中間評価（2 年目）の評点に関しては群間で差がみられなかった。しかし開催通知や P0 の出席の「後」の事後評価の評点に関しては「班会議の開催通知があり、P0 が出席した」研究課題の行政点、評価点が高かった。

表 17-1 ~ 表 17-4 に、全ての研究課題について、事前評価、中間評価（1 年目）、中間評価（2 年目）、事後評価の、専門点、行政点、評価点の間の相関係数を示した。

表 17-1 に、事前評価との相関係数を示した。事前評価においては、専門点と行政点の間の相関はみられなかったが、評価点と専門点及び行政点の間の正の相関がみられた。専門点に関しては事前評価と事後評価の間の正の相関がみられた。行政点に関しては事前評価と中間評価（1 年目）の間の正の相関がみられた。評価点に関しては事前評価と中間評価（1 年目）、事前評価と事後評価の間の正の相関がみられた。また事前評価の専門点と事後評価の行政点及び評価点の間、事前評価の行政点と中間評価（1 年目）の評価点の間、事前評価の評価点と中間評価（1 年目）の専門点及び行政点、事後評価の専門点及び行政点の間の正の相関がみられた。

表 17-2 に、中間評価（1 年目）との相関係数を示した。中間評価（1 年目）においては、専門点と行政点の間の相関はみられなかったが、評価点と専門点及び行政点の間の正の相関がみられた。専門点、評価点に関しては、中間評価（1 年目）と中間評価（2 年目）及び事後評価の間の正の相関がみられた。行政点に関しては中間評価（1 年目）と事後評

価の間の正の相関がみられた。また中間評価（1年目）の専門点と中間評価（2年目）及び事後評価の評価点の間、中間評価（1年目）の評価点と中間評価（2年目）の専門点及び事後評価の行政点の間の正の相関がみられた。

表 17-3 に、中間評価（2年目）との相関係数を示した。中間評価（2年目）においては、専門点と行政点の間の相関はみられなかったが、評価点と専門点及び行政点の間の正の相関がみられた。専門点、行政点、評価点ともに、中間評価（2年目）と事後評価の間の正の相関がみられた。また中間評価（2年目）の行政点と事後評価の専門点及び評価点の間、中間評価（2年目）の評価点と事後評価の専門点及び行政点の間の正の相関がみられた。

表 17-4 に、事後評価における評点間の相関係数を示した。専門点と行政点の間の正の相関がみられ、評価点と専門点及び行政点の間の正の相関がみられた。

3. 研究成果の状況

表 18 に、研究分野別にみた、原著論文（和文）の発表件数を示した。全体では、平均値が 3.45 件、中央値は 1.00 件で、多い方にすそをひく分布であった。発表件数 0 件が約 4 割、4 件以下が約 8 割であった。研究分野別にみると、平均値では「地域」が低く、テロが高かったが、ノンパラメトリック検定では統計的に有意な差はみられなかった。「テロ」に関しては 39 件と極端に多い研究があることが影響していると考えられる。

表 19 に、研究分野別にみた、原著論文（英文等）の発表件数を示した。全体では、平均値が 4.98 件、中央値は 1.00 件で、多い方にすそをひく分布であった。発表件数 0 件が約 5 割、4 件以下が約 8 割であった。研究分野別では「地域」、「テロ」の件数が少なかった。

表 20 に、研究分野別にみた、その他論文（和文）の発表件数を示した。全体では、平均値

が 4.72 件、中央値は 1.00 件で、多い方にすそをひく分布であった。発表件数 0 件が約 4 割、6 件以下が約 8 割であった。研究分野別では、「生活衛生」の件数が多かったが、統計的に有意な差はみられなかった。

表 21 に、研究分野別にみた、その他論文（英文等）の発表件数を示した。全体では、平均値が 0.40 件、中央値は 0.00 件で、発表件数 0 件が 8 割であった。研究分野別では、「地域」、「テロ」の件数が少なかった。

表 22 に、研究分野別にみた、国内学会での発表件数を示した。全体では、平均値が 13.28 件、中央値は 7.00 件で、多い方にすそをひく分布であった。発表件数 4 件以下が約 3 割、19 件以下が約 8 割であった。研究分野別では、「地域」、「テロ」の件数が少なかった。なお「テロ」に関しては平均値が高かったが、これは 90 件と極端に多い研究があることが影響していると考えられる。

表 23 に、研究分野別にみた、国際学会等での発表件数を示した。全体では、平均値が 3.68 件、中央値は 2.00 件で、多い方にすそをひく分布であった。発表件数 0 件が約 4 割、6 件以下が約 8 割であった。研究分野別では、「地域」、「テロ」の件数が少なかった。

表 24 に、研究分野別にみた、特許の出願及び取得の件数を示した。9 割以上の研究課題が特許の出願・取得を実施していなかった。

表 25 に、研究分野別にみた、施策への反映の件数を示した。全体では、平均値が 1.22 件、中央値は 0.00 件で、多い方にすそをひく分布であった。施策に反映した件数 0 件が約 7 割、2 件以下が約 9 割であった。研究分野で件数の差はみられなかった。

表 26 に、研究分野別にみた、普及・啓発活動の件数を示した。全体では、平均値が 1.85 件、中央値は 0.50 件で、多い方にすそをひく分布であった。普及・啓発活動の件数 0 件が 5 割、2 件以下が約 8 割であった。研究分野で件数の差はみられなかった。

4. 研究成果と評点との関連

表 27-1、表 27-2 に、研究成果及び研究費総額との順位相関係数を示した。原著論文（和文）の発表件数に関しては、原著論文（英文等）の発表件数、その他論文（英文等）の発表件数、国内学会での発表件数、国際学会等での発表件数との正の相関がみられた。また原著論文（英文等）の発表件数に関しては、原著論文（和文）の発表件数、その他論文（英文等）の発表件数、国内学会での発表件数、国際学会等での発表件数との正の相関がみられた。

その他論文（和文）の発表件数に関しては、その他論文（英文等）の発表件数、国内学会での発表件数、国際学会等での発表件数、施策への反映の件数との正の相関がみられた。またその他論文（英文等）の発表件数に関しては、原著論文（和文）の発表件数、原著論文（英文等）の発表件数、その他論文（和文）の発表件数、国内学会での発表件数、国際学会等での発表件数との正の相関がみられた。

国内学会での発表件数に関しては、原著論文（和文）の発表件数、原著論文（英文等）の発表件数、その他論文（和文）の発表件数、その他論文（英文等）の発表件数、国際学会等での発表件数との正の相関がみられた。また国際学会等での発表件数に関しては、原著論文（和文）の発表件数、原著論文（英文等）の発表件数、その他論文（和文）の発表件数、その他論文（英文等）の発表件数、国内学会での発表件数との正の相関がみられた。

施策への反映の件数に関しては、その他論文（和文）の発表件数、普及・啓発活動の件数との正の相関がみられた。また普及・啓発活動の件数に関しては、施策への反映の件数との正の相関がみられた。

研究費総額に関しては、原著論文（和文）の発表件数、原著論文（英文等）の発表件数、その他論文（英文等）の発表件数、国内学会

での発表件数、国際学会等での発表件数との正の相関がみられた。しかし、施策への反映の件数、普及・啓発活動の件数との相関はみられなかった

表 28 に、原著論文（和文）の発表の有無別にみた評点の平均値と標準偏差を示した。発表の有無で、事前評価、中間評価の評点の差はみられなかったが、事後評価においては、発表している研究課題の方が専門点、行政点、評価点ともに高かった。

表 29 に、原著論文（英文等）の発表の有無別にみた評点の平均値と標準偏差を示した。発表の有無で、事前評価、中間評価の評点の差はみられなかったが、事後評価においては、発表している研究課題の方が評価点が高かった。

表 30 に、その他論文（和文）の発表の有無別にみた評点の平均値と標準偏差を示した。発表している研究課題の方が、事前評価の専門点が低く、中間評価（2年目）の専門点及び評価点が高かった。

表 31 に、その他論文（英文等）の発表の有無別にみた評点の平均値と標準偏差を示した。発表の有無で、事前評価、中間評価（1年目）の評点の差はみられなかった。中間評価（2年目）においては、発表している研究課題の方が行政点、評価点が高く、事後評価においては、発表している研究課題の方が専門点、行政点、評価点ともに高かった。

表 32 に、国際学会等での発表の有無別にみた評点の平均値と標準偏差を示した。発表の有無で、事前評価、中間評価の評点の差はみられなかったが、事後評価においては、発表している研究課題の方が行政点、評価点が高かった。

表 33 に、施策への反映の有無別にみた評点の平均値と標準偏差を示した。施策に反映した研究課題の方が事前評価における専門点、評価点が低かった。また反映の有無で、中間評価、事後評価の評点の差はみられなかった。

表 34 に、普及・啓発活動の有無別にみた評点の平均値と標準偏差を示した。普及・啓発活動を実施した研究課題の方が事前評価、中間評価（1年目）の評価点が低かった。また活動の有無で、中間評価（2年目）、事後評価の評点の差はみられなかった。

表 35-1、表 35-2 に、研究成果と評点の間の順位相関係数を示した。原著論文（和文）の発表件数、原著論文（英文等）の発表件数に関しては、事前評価、中間評価の評点との相関はみられなかったが、事後評価の専門点、行政点、評価点との正の相関がみられた。

その他論文（和文）の発表件数に関しては、事前評価、中間評価（1年目）、事後評価の評点との相関はみられなかったが、中間評価（2年目）の専門点及び評価点との正の相関がみられた。

その他論文（英文等）の発表件数に関しては、事前評価、中間評価（1年目）の評点との相関はみられなかったが、中間評価（2年目）の行政点及び評価点、事後評価の専門点、行政点、評価点との正の相関がみられた。

国内学会での発表件数に関しては、事前評価の専門点及び評価点、中間評価（2年目）の行政点、事後評価の専門点、行政点、評価点との正の相関がみられた。

国際学会等での発表件数に関しては、事前評価、中間評価（1年目）の評点との相関はみられなかったが、中間評価（2年目）の行政点、事後評価の専門点、行政点、評価点との正の相関がみられた。

施策への反映の件数に関しては、事前評価、中間評価（1年目）、事後評価の評点との相関はみられなかったが、中間評価（2年目）の専門点との正の相関がみられた。

普及・啓発活動の件数に関しては、事前評価、中間評価、事後評価のいずれの評点との相関もみられなかった。

D. 考察

1. 事前・中間・事後評価の変化と相関

専門点は事前評価と比較して事後評価が低く、行政点と評価点は事前評価と比較して中間評価、事後評価が低かった。事前評価と中間・事後評価はそれぞれ異なる委員によって行われているため厳密には比較できないが、事前評価の段階で期待されていた成果が十分に得られていないことを示唆している。したがって、全ての研究課題に対して、事前評価の段階で期待されていた学術的・行政的な成果を十分に認識した上で研究を遂行できるように支援する必要がある。

専門点の事前・中間・事後の相関に関しては、事前評価と事後評価の間、中間評価（1年目）と中間評価（2年目）の間、中間評価（1年目）と事後評価の間、中間評価（2年目）と事後評価の間で正の相関がみられた。これは、最初の研究計画が優れたものでないと、最終的に学術的に優れた研究成果を得ることが困難であることを示唆している。また学術的観点からの評価は中間評価から事後評価まで比較的一貫していることを示唆している。したがって、事前評価や中間評価の低い研究課題に対して、特に学術的側面から、研究期間全体を通じて優れた成果を得られるように支援する必要がある。

行政点の事前・中間・事後の相関に関しては、事前評価と事後評価の間の相関はみられなかったが、事前評価と中間評価（1年目）の間、中間評価（1年目）と事後評価の間、中間評価（2年目）と事後評価の間といった、直近の評価の間で正の相関がみられた。これは、当初の行政ニーズに十分に適合した研究計画を策定したとしても、その後のニーズの変化に対応していかなければ、最終的に行政的に有用な成果を得られないことを示唆している。また直近の評価の間の関連が強いことから、比較的短期間に求められる行政ニーズに対応している研究課題の方が行政的観点か

らの評価が高くなることが示唆される。したがって行政点の低い研究課題に対して、研究期間の途中段階であっても、行政ニーズに応じた適時の成果を比較的短期間に得られるように支援する必要がある。

評価点の事前・中間・事後の相関に関しては、事前評価と中間評価（1年目）の間、事前評価と事後評価の間、中間評価（1年目）と中間評価（2年目）の間、中間評価（1年目）と事後評価の間、中間評価（2年目）と事後評価の間で正の相関がみられた。この結果は上述した専門点の傾向と行政点の傾向を反映したものと考えられる。

専門点と行政点の間の相関に関しては、事前評価、中間評価においてはみられなかったが、事後評価において正の相関がみられた。研究課題には学術面に重点が置かれているものと行政面に重点が置かれているものがあり、研究計画の段階（事前評価）や研究進捗の段階（中間評価）においては両側面の評価が必ずしも一致しないと考えられる。しかし研究期間を通じて、学術面と行政面の両面からの評価と改善に向けた助言等を受けることによって、最終的な研究成果の段階（事後評価）では両側面の評価が一致したと考えられる。

研究分野別では、中間評価では分野間の評点の差はみられなかったが、「地域」に関しては、事前評価の専門点、行政点、評価点、事後評価の行政点と評価点が低かった。したがって「地域」の研究課題に対しては、他の研究分野以上に積極的に働きかけ、学術面、行政面の両面の研究成果が得られるように支援する必要がある。

2. 研究課題の進捗状況

本研究事業では、平成19年度より、研究班会議の開催をP0に通知してもらい、P0は可能な限り出席して、研究遂行に必要な情報を提供するという活動を実施してきた。

研究班会議の開催通知があった研究課題の割合は、研究開始1年目で84.6%、2年目で56.8%、3年目で70.4%、P0が出席した研究課題の割合は、1年目で57.7%、2年目で37.8%、3年目で59.3%と、いずれも2年目に減少し、3年目に増加していた。研究終了年度の3年目は研究期間が残り少ないためP0が支援できる余地は少ないと考えられる。それに対して1年目、2年目は、研究成果の向上のために研究の方向性や内容等を改善することが可能であり、P0が支援できる部分も比較的大きいと考えられる。したがって、研究開始1年目、2年目、特に研究開始当初の積極的な意識が若干低下すると考えられる2年目の研究課題に対して積極的な支援を実施する必要がある。また研究期間が2年間の課題も多いことから、2年目の早い段階での支援が重要になると考えられる。

研究課題の進捗状況と評点との関連では、開催通知の有無、P0の出席の有無で事前評価、中間評価の評点の差はみられなかった。しかし事後評価に関しては、開催通知やP0の出席の「前」の中間評価（2年目）の評点の差はみられなかったが、開催通知やP0の出席の「後」の事後評価で、「班会議の開催通知があり、P0が出席した」研究課題の行政点、評価点が高かった。したがって今後はP0が班会議に出席できなかった研究課題に対する支援を強化する必要がある。具体的には、研究代表者に対して電話や訪問などを行い、研究の進捗状況の把握や研究の推進に有用な情報提供などを実施する必要がある。

3. 研究成果の状況

本研究で使用した研究成果の変数は、学術的成果（原著論文（和文）の発表件数、原著論文（英文等）の発表件数、その他論文（和文）の発表件数、その他論文（英文等）の発表件数、国内学会での発表件数、国際学会等での発表件数）と行政的成果（特許の出願・

取得、施策への反映、普及・啓発活動)に大別することができる。

学術的成果に関しては、「地域」、「テロ」で英文等による原著論文やその他論文の発表件数、国内・国際学会での発表件数が少なかった。これらの研究分野ではわが国に特有の行政課題に取り組んでいる研究が多く、国際的に普遍的な知見を国際雑誌や国際学会に発表することが困難である場合もある。しかしわが国の行政課題には諸外国にも共通する部分も多く、研究で得られた行政課題の解決策やわが国の優れた特徴を世界に向けて積極的に情報発信していくことによって、わが国及び諸外国の健康安全・危機管理システムの改善に貢献できると考えられる。

行政的成果に関しては、特許の出願・取得が0件の課題が9割以上、施策への反映が0件の課題が約7割、2件以下が約9割、普及・啓発活動が0件の課題が約5割、2件以下が約8割と、非常に少なかった。本研究事業は行政ニーズに対応するための研究課題が多いにも関わらず、このような行政的成果の少なさは重大な問題であると考えられる。行政点の事前・中間・事後の相関でも考察したように、本研究事業の研究課題は比較的短期間に求められる行政ニーズに対応していく必要がある、そのためには行政担当者との密接にコミュニケーションを図り、研究テーマに関連する様々な行政課題に柔軟に対応して、施策に反映できる多くの研究成果を産出していく必要がある。

研究成果の間の関連については、学術的成果の変数(論文発表、学会発表の件数)の間で互いに正の相関がみられ、行政的成果の変数(施策への反映、普及・啓発活動)の間で互いに正の相関がみられたが、学術的成果と行政的成果との相関はほとんどみられなかった。これは、学術面の成果と行政面の成果が独立して産出されていることを示しており、学術的に新しい知見だけでなく行政課題の解

決に向けた知見の両方を追求する厚生労働科学研究においては両面からの評価を今後も継続していく必要があることを示唆している。しかし一方で、学術的成果と行政的成果は両立しないというわけではなく、学術的な指向の強い課題では行政的側面、行政的な指向の強い課題では学術的側面と、それぞれの弱い側面を強化して両方の成果を効果的に産出していく必要がある。

研究費と研究成果との関連については、研究費総額と学術的成果との正の相関がみられたが、行政的成果との相関はみられなかった。研究費が高額な研究課題は研究班の規模(研究分担者、研究協力者の数等)が大きく、1人当たりの学術的成果(論文の投稿、学会での発表など)が一定であれば、学術的成果の総数は多くなると考えられる。一方、研究課題に対応する行政課題は研究費の多少に関わらず一定であるため、行政的成果も研究費に関係なくほとんど一定になると考えられる。

4. 研究成果と評点との関連にみる研究評価の妥当性

① 研究評価の妥当性の基本的な考え方

研究評価システムが適正に運営されているかどうかを「評価」するためには、研究評価が研究成果と結びついているかどうかを検証する必要がある。具体的には、事前評価においてより多くの研究成果を産出できる研究計画を高く評価しているか、中間評価においてより多くの研究成果を産出している、あるいは産出すると期待される研究の進捗状況を高く評価しているか、そして事後評価において産出された研究成果を高く評価しているか、つまり研究成果に対して研究評価が「妥当」であるかどうかを検証する必要がある。

このような妥当性は「基準妥当性(criterion validity)」、つまり基準(gold standard)との相関の程度に相当する。基準妥当性は「同時妥当性(concurrent

validity)」、つまり「同時に」測定される基準との相関と、「予測妥当性 (predictive validity)」、つまり「将来」測定される(現在は測定できない)基準との相関の2つに分類することができる。研究評価に当てはめると、事前評価と研究成果(将来測定される基準)との相関は予測妥当性、事後評価と研究成果(同時に測定される基準)との相関は同時妥当性に相当する。また中間評価は予測妥当性と同時妥当性の両方の要素をもつが、研究年次が少ないほど予測妥当性、研究年次が多いほど同時妥当性の性質が強くなると考えられる。

②評点全体の基準妥当性の検証

学術的成果に関しては、事前評価は国内学会での発表以外の学術的成果の変数との相関がみられなかった。中間評価(1年目)は全ての学術的成果の変数との相関がみられなかった。中間評価(2年目)は、原著論文(和文、英文等)との相関はみられなかったが、その他論文(和文、英文等)、国内・国際学会での発表との正の相関がみられた。そして事後評価はその他論文(和文)以外の学術的成果の変数との正の相関がみられた。以上の結果から、事前評価は学術的成果をほとんど予測できていないこと、事後評価は学術的成果を十分に反映していることが示された。また中間評価に関しては、予測妥当性の側面が強い1年目では学術的成果をほとんど予測できていないこと、同時妥当性の側面が強い2年目では学術的成果をある程度反映していることが示された。

行政的成果に関しては、事前評価は、施策に反映した研究課題、普及・啓発活動を実施した研究課題の方がむしろ低かった。中間評価はほとんどの行政的成果の変数との相関がみられなかった。事後評価は全ての行政的成果の変数との相関がみられなかった。以上の結果から、事前評価は行政的成果をほとんど

予測できていないこと、事後評価も行政的成果をほとんど反映していないこと、中間評価も同様の傾向であることが示された。

研究評価(事後評価、中間評価)の同時妥当性に関しては、学術的成果に対しては高かったが、行政的成果に対しては低かった。事後評価の対象となる研究報告書や中間評価の対象となる研究計画書(継続申請用)にはこれまでの成果を記載する項目があり、そこに記載された論文発表や学会発表の数や内容が評価の参考にされるため、評点と学術的成果との相関がみられたと考えられる。しかし行政的成果に関しては、研究報告書には知的財産権の出願・登録状況を記載する項目はあるが、施策への反映や普及・啓発活動を記載する項目が設定されていない。また研究計画書(継続申請用)には成果を記載する項目に「研究課題の実施を通じた政策提言(寄与した指針又はガイドライン等)」を記載するように指示されているが、論文発表のみが記載され行政的成果が記載されていない研究計画書がほとんどである。そのため、評価にあたっての判断材料が十分でなく、行政的成果が評点に反映されにくいと考えられる。したがって、研究報告書や研究計画書の様式に行政的成果を記載する項目を設定するとともに、それらを具体的かつ詳細に記載するように研究代表者に周知徹底する必要がある。

研究評価(事前評価、中間評価)の予測妥当性に関しては、学術的成果、行政的成果ともに低かった。これは一つには、事前評価や中間評価の対象となる研究計画書の記載事項に将来期待される研究成果を予測できる内容が十分に記載されていないことが考えられる。記載事項には「期待される成果」が含まれているが、健康への寄与、行政施策への貢献といった抽象的な記述がほとんどであり、例えば、インパクトファクターの高い学術雑誌に掲載されるレベルの知見を得られるのか、どの法律の改正、どの審議会・検討会に対して