

201504012A

厚生労働科学研究費補助金

特別研究事業

厚生労働科学研究における研究成果の
アウトカム評価の指標及び
手法の開発に関する研究

平成 27 年度 研究報告書

研究代表者 尾島 俊之

(浜松医科大学健康社会医学講座教授)

厚生労働科学研究における研究成果のアウトカム評価の指標及び手法の開発に関する研究 研究結果の概要

【目的】

厚生労働科学研究における研究成果のアウトカムを適切に把握するための指標及び手法を提案し、研究成果を総合的かつ長期的に評価する仕組みを検討することを目的とした。

【方法】

米国におけるアウトカム評価の調査、行政的アウトカムの評価方法の検討、アウトカム評価の活用方策の検討、地方自治体の視点からの検討、学術的アウトカムの評価方法の検討、経済効果に関する検討を行い、最後にそれらの結果をふまえてアウトカム評価の指標・手法の検討を行った。

【結果と考察】

推奨されるアウトカム指標としては、(1) 死亡率（米国 NIH における主要なアウトカム評価指標となっており、国内においても肝炎に関する研究が肝臓死亡率の減少に寄与している可能性が示唆されているなどの事例が存在）、(2) 研究領域に特有の指標（例えば、移植成績、患者に適切に説明する医療機関の割合、自殺企図の再発率など）、(3) 制度形成に関する定性的評価（例えば、支援制度、健診制度、診療報酬改定などに資する研究成果など）が挙げられる。

一方で、現時点で推奨を保留とする指標としては、QALY (Quality Adjusted Life Year、本質的に健康に関する最も重要な指標のひとつであるが、研究の評価指標として未確立)、金銭表示での便益（研究評価としての実例が希）、学術論文の被引用数（学術面でのアウトカム評価指標として挙げられることが多いが、系統的に把握するための費用や労力・研究領域等による格差・行政的な成果が軽視されるおそれ等が課題）、マスメディアへの掲載数（社会へのインパクトを定量化する指標となりうるが、定量化の方法が確立していない・研究の成果以外の要因が大きいなどが課題）がある。

アウトカム評価の方法として、制度形成への貢献などについては、引き続き研究者から、また必要に応じて行政担当者からの報告を求める意義は大きいと考えられる。一方で、諸外国や多分野における実施状況を注視しながら、過度な負担にならないよう、データベースの検索や統計資料の解析などを中心にアウトカム評価を実施することが推奨される。

良いアウトカムが得られる研究を活性化するためには、(1) 目的・ゴールの明確な設定、(2) 行政的な成果の推進、(3) 学術的な成果の推進、(4) 社会への情報発信の推進、(5) 体制の構築が推奨される。

【結論】

厚生労働科学研究がより効果的に実施され、国民の健康等にさらに一層貢献できるように、期待される研究成果を明確化して、それに沿ったアウトカム評価を推進していくことが必要である。

目 次

| | |
|---|----|
| I. 総括研究報告書 | 1 |
| 尾島 俊之（浜松医科大学健康社会医学） | |
| 参考資料 1 米国におけるアウトカム評価に関するインタビュー記録 | 8 |
| 参考資料 2 NIH 研究の効果 | 28 |
| II. 分担研究報告書 | |
| 1. 厚生労働科学研究における研究成果の行政的アウトカムに関する評価方法の検討 ---- | 37 |
| 神里 彩子（東京大学医科学研究所） | |
| 2. 「自殺対策のための戦略研究」の施策への活用プロセスに関する分析 | 51 |
| 福井 次矢・萱間 真美（聖路加国際大学） | |
| 3. 地方自治体の視点からの厚生労働科学研究における研究成果のアウトカム評価の指標 及び手技に関する研究 | 71 |
| 前田 秀雄（東京都医学総合研究所） | |
| 4. 研究成果の学術的アウトカムの評価方法の検討 | 75 |
| 緒方 裕光（国立保健医療科学院研究情報支援研究センター） | |
| 5. 厚生労働科学研究成果の経済効果に関する検討 | 91 |
| 満武 巨裕（医療経済研究機構） | |
| III. 研究成果の刊行に関する一覧 | 95 |

厚生労働科学研究における研究成果のアウトカム評価の 指標及び手法の開発に関する研究

研究代表者 尾島 俊之（浜松医科大学医学部健康社会医学講座 教授）

研究要旨

厚生労働科学研究における研究成果のアウトカムを適切に把握するための指標及び手法を提案し、研究成果を総合的かつ長期的に評価する仕組みを検討することを目的とした。米国におけるアウトカム評価の調査、行政的アウトカムの評価方法の検討、アウトカム評価の活用方策の検討、地方自治体の視点からの検討、学術的アウトカムの評価方法の検討、経済効果に関する検討を行い、最後にそれらの結果をふまえてアウトカム評価の指標・手法の検討を行った。推奨されるアウトカム指標としては、死亡率、研究領域に特有の指標、制度形成に関する定性的評価が挙げられる。一方で、現時点で推奨を保留とする指標としては、QALY、金銭表示での便益、学術論文の被引用数、マスメディアへの掲載数がある。厚生労働科学研究がより効果的に実施され、国民の健康等にさらに一層貢献できるように、研究の目標と期待される成果を明確化して、それに沿ったアウトカム評価を推進していくことが必要である。

研究分担者

緒方 裕光 国立保健医療科学院研究情報
支援研究センターセンター長
前田 秀雄 東京都医学総合研究所理事長
神里 彩子 東京大学医科学研究所公共政
策研究分野研究倫理支援室特任准教授
福井 次矢 学校法人聖路加国際大学・聖路
加国際病院院長
萱間 真美 聖路加国際大学大学院精神看
護学教授
満武 巨裕 一般財団法人医療経済研究・社
会保険福祉協会 医療経済研究機構研
究副部長

研究協力者

角田 秋 聖路加国際大学大学院精神看護
学准教授

野添 篤毅 愛知淑徳大学名誉教授・国立保
健医療科学院客員研究員
水島 洋 国立保健医療科学院研究情報支
援研究センター上席主任研究官
佐藤 洋子 国立保健医療科学院研究情報
支援研究センター研究員
佐々木 康則（元）一般財団法人医療経済
研究・社会保険福祉協会 医療経済研
究機構 企画渉外部部長

A. 研究目的

厚生労働科学研究では、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」
を実施している。

厚生労働科学研究における研究課題評価で

は、事前評価（研究成果を達成する潜在能力があるか）、中間評価（研究成果の達成に向けて進捗しているか）、事後評価（研究成果を達成したか）が実施されている。しかし事後評価においては、論文発表、学会発表、特許、施策への反映等の「アウトプット」の評価にとどまっており、それらが研究成果の利用者にどのように貢献したか、学術的、行政的にどのようなインパクトを及ぼしたか、国民の健康・安全を向上させたか、といった「アウトカム」が十分に評価されていない。そのため、研究成果を総合的、長期的に評価するための指標及び手法を確立させるための研究が不可欠である。なお、国の研究開発評価に関する大綱的指針（平成24年内閣総理大臣決定）によると、アウトカム指標とは、「成果の本質的又は内容的側面であり、活動の意図した結果として、定量的又は定性的に評価できる、目標の達成度を測る指標」と定義されている。

そこで本研究では、厚生労働科学研究における適切なアウトカム指標及び評価方法を提案し、研究成果を総合的かつ長期的に評価する仕組みを検討することを目的とした。

B. 研究方法

1. 米国におけるアウトカム評価の調査

米国国立衛生研究所（NIH, National Institute of Health）科学審査センター（CSR, Center for Scientific Review）・外部委託研究局（OER, Office of Extramural Research）及び米国保健医療向上研究機構（AHRQ, Agency for HealthCare Research and Quality）において、助成したまたは自ら実施した研究プロジェクトの研究成果の評価方法、研究成果のアウトカムの評価に関する取り組みなどに関して現地訪問調査を実施した。

2. 行政的アウトカムの評価方法の検討

(1) 行政担当者を対象とした自記式調査票調

査

厚生労働科学研究の全研究事業（27事業）についての厚生労働省内の各研究事業担当者を対象として、2015年11月に厚生労働省厚生科学課からのメールによって自記式調査票調査を実施した。調査票として、各研究事業全体における研究のアウトカム評価に関しての「研究事業シート」と、行政的観点から成果のあったと考える（法改正、制度改正、通知につながったなど）グッドプラクティス研究課題の提示と、それについての情報を得ることを主目的とする「研究課題シート」の2種類について調査した。

(2) インタビュー調査（面接調査）

研究事業数が最も多い「疾病・障害対策研究分野」の中から「肝炎等克服政策研究事業」、また「健康安全確保総合研究分野」の中から「食品の安全確保推進研究事業」選択し、行政担当者に、研究成果が施策（各種基準、ガイドライン等）に反映された経緯や背景、研究成果から施策への反映の具体的な流れなどについてインタビュー調査を行った。また、肝炎等克服政策研究事業については、研究代表者にもインタビュー調査を行った。

(3) 文献調査

調査対象とする研究課題の成果報告書、当該研究課題の成果が厚生労働省の審議会等で報告された際の資料、研究成果が行政施策として発出された際の資料等の調査を行った。

3. アウトカム評価の活用方策の検討

「自殺対策のための戦略研究」の2つの研究プロジェクト（ACTION-J, NOCOMIT-J）の概要を文献検討し、うちACTION-Jの統括研究責任者、自殺対策担当部門の行政担当者（厚生労働技官）、戦略研究で本研究の評価を担当した外部評価者の各1名を対象に、半構造化インタビューを実施した。インタビュー内容の逐

語録を作成、本戦略研究のアウトプット、アウトカム（政策・社会への還元まで）について、どのような視点でとらえているか、研究の内容に関してどのように評価しているかに関連する内容を抽出した。

4. 地方自治体の視点からの検討

地方自治体等における政策決定及び過程を整理し、特に参酌すべき段階を中心に、研究成果の活用という視点で、厚生労働科学研究のアウトカム評価の指標を検討した。

5. 学術的アウトカムの評価方法の検討

(1) 論文数および論文の被引用件数等による定量的評価

厚生労働科学研究成果データベースに収載されている研究成果論文のデータを用いて、学術情報データベース「Web of Science®」

（トムソン・ロイターにより提供）を用いて、定量的評価を行った。具体的には、論文数、被引用数について、研究課題別、研究事業別、研究代表者別、論文の著者機関別、論文の発表年別に集計を行った。

(2) 行政効果報告に含まれる文章情報に基づく評価

平成26年度厚生労働科学研究成果データベースの行政効果報告（公開件数：626件）の「専門的・学術的観点からの成果」、「臨床的観点からの成果」、「その他行政的観点からの成果」、「ガイドライン等の開発」、「その他のインパクト」の文章データについてテキストマイニング分析を行った。分析は、形態素解析（文章データを単語ごとに分解）、キーワード抽出（出現頻度の高い単語（キーワード）を抽出）、固有表現抽出（予め指定した成果報告に出現する特有の表現にマッチしたキーワードおよびキーフレーズを抽出し、出現回数の集計）、クラスタリング分析（キーフレーズの評点を「低い」～「かなり高い」の5段階にカ

テゴライズし、各キーフレーズの評点×出現回数を加算）という手順で実施した。

6. 経済効果に関する検討

厚生労働科学研究データベースの新聞記事に関する入力情報と新聞広告料金表等をもとにして、各研究等の新聞記事による情報発信について金額に換算したアウトカムの推計を試みた。

7. アウトカム評価の指標・手法の検討

以上の結果を踏まえて、研究成果のアウトカム指標について検討し、推奨する指標、現時点で推奨を保留する指標を選別した。また、アウトカム評価の方法、良いアウトカムが得られる研究とするためのポイントを検討した。

C. 研究結果

1. 米国におけるアウトカム評価の調査

米国におけるアウトカム評価に関するインタビュー記録を参考資料1に、NIH研究の効果としてホームページで一般に公表されている資料の日本語訳を参考資料2に示す。NIHにおいては、健康、社会、知識という3つの視点でアウトカム評価を行っている。健康についての指標としては平均寿命や死亡率が最重要の指標として用いられている。また、社会については、経済成長への貢献、知識については、知識ベースの構築や世界的な科学者の育成などが評価項目として使用されている。AHRQにおいては、一般紙への掲載、米国議会の法案提案書での引用、経済的評価、公平性などについて評価指標として議論がされているが、まだ検討が始まったばかりであり、実用化には至っていない。

2. 行政的アウトカムの評価方法の検討

研究成果の行政的アウトカムが得られるためには、研究成果が行政施策に結びつくこと

が必須である。それには行政側、研究者側それぞれでやらなければならないことがある。研究終了後一定期間が経過した後に、行政施策の効果について評価・検証する機会を設定することが重要且つ必要である。研究成果の行政的アウトカムを評価するために評価指標を持っている研究事業は少なく、また、十分に検討されていなかった。今後、アウトカム評価を各事業で実施していくにあたって、各事業において何を指標とするべきかを検討する必要があるが、評価指標は客観的評価ができる(数値化された)指標でなければならず、その知見の蓄積が当該研究事業の評価軸として確立していくことが期待される。また、アウトカム評価を実施することにより、各研究事業において、研究の行政的効果・アウトカムを意識して研究課題設定を検討し、適切な課題設定へとつながることも期待される。このような研究事業におけるアウトカム評価に加えて、継続的な研究を行っている研究班においては、研究計画の一部にそれを組み入れること推奨することも一案として考えられる。

3. アウトカム評価の活用方策の検討

「自殺対策のための戦略研究」は2つの研究プロジェクト(ACTION-J, NOCOMIT-J)からなり、通常の厚生科学研究の評価の枠組みとは異なる枠組みで行われている。研究のゴールについての認識には研究責任者、行政担当者、外部評価者の3者間で相違がみられた。自殺者数の減少は本研究の直接的な効果とはいえないが、本研究は、自殺未遂者対策に関する政策立案プロセスへの貢献に一定の評価が可能であろう。海外の先行研究では自殺未遂者対策による長期的な自殺防止効果は限定的で、むしろ医療経済的な貢献があるとされている。本研究の長期的貢献については、なお継続した評価が必要である。

4. 地方自治体の視点からの検討

地方自治体の視点での重要なアウトカム評

価指標として、(1) 施策方針への反映(当該分野の重要性、地域社会の高いニーズ、効果的な実施方法を示唆する研究)、(2) 具体的事業への反映(具体的事業の技術の妥当性、効率的運用、水準保持の基盤となる科学的知見、地域の実情を科学的に反映するための根拠となる研究)、(3) 住民に対する普及啓発への活用(科学的根拠についても言及した正確な情報を提供)、(4) 関係機関との連携への活用(科学的根拠に基づいた関係機関との理解の共有)が抽出された。

5. 学術的アウトカムの評価方法の検討

厚生労働科学研究成果データベースに登録された書誌情報を用いた分析から、課題別、事業別、年度別、機関別などの定量的評価が可能であることが分かった。また、同データベースに登録された文章情報に関するテキストマイニングの手法により、ある程度は客観的な評価が可能であることが示された。したがって、例えば「成果の高い研究を抽出する」、「継続課題設定の参考にする」などの目的で活用することは可能である。また、文章情報の評価への活用においては、実際の専門家による評価との比較を通じて、キーワードやキーワードへの重みづけなどを検討していく必要がある。

6. 経済効果に関する検討

厚生労働科学研究の成果が新聞記事として取り上げられて国民に発信されたものについて、広告料金表により、金銭価値として評価を行うことができた。ただし、記事の全てが研究成果ではない場合がある、複数年で行われている研究の場合どの年度の研究成果であるのか特定することが困難である、「その他のインパクト」の入力が標準化されていない等の理由で記事を特定することが困難な場合があるなどの問題点があり、現時点においてこの方法による経済的な評価をアウトカム評

価として採用することは妥当ではない。今後、厚生労働科学研究成果データベースの「その他のインパクト」の入力方法の標準化、プレスリリース等の活動の推奨も有用であろう。

D. 考察

1. アウトカム評価指標

(1) 推奨する指標

・死亡率

推奨する根拠としては、米国NIHにおける主要なアウトカム評価指標となっていること、肝炎に関する研究が肝臓死亡率の減少に寄与している可能性が示唆されているなどの事例が存在していることが挙げられる。一方で、留意点として、死亡率以外のアウトカムを目指す研究も多いことから、そのような研究においては使用することができない。また、最終的に死亡率の減少を目指す研究であっても、死亡率の減少効果が現れるまでには長期間を要する。また、研究と研究以外の寄与分を明確に分離することは困難である。これらの点に留意しながら使用する必要がある。

・研究領域に特有の指標

例えば、移植成績、患者に適切に説明する医療機関の割合、自殺企図の再発率など、客観的評価ができる(数値化された)指標がある。推奨する根拠としては、実際の研究事業で活用または提唱されている事例があることが挙げられる。一方で、留意点としては、適切な指標がみあたらない研究領域も多い点がある。

・制度形成に関する定性的評価

例えば、支援制度、健診制度、診療報酬改訂に資する研究成果などがある。推奨する根拠としては、その研究事業の本質的な目的に合致することが多いこと、実際に制度形成に結実している研究は多いことが挙げられる。一方で、留意点として、研究実施から制度形成までタイムラグがある場合などにおいて、制

度形成に結実したことの把握が困難なことがある点が挙げられる。

(2) 現時点で推奨を保留する指標

・QALY (Quality Adjusted Life Year、質調整生存年数)

本質的に健康に関する最も重要な指標の一つである。また医療技術の評価指標としての活用はある。しかしながら、研究の評価指標としては確立していない。また、米国AHRQにおけるプロジェクト評価指標としての検討が始まっているが、まだ実用化の目処は立っていない。

・金銭表示での便益

財政的に厚生労働科学研究の充実を図ろうとした場合に、費用対効果を示すことができると説得力が増す。しかしながら、OECD諸国等の諸外国においても、研究に関する費用便益分析の実例は希であり、その評価方法は確立していない。

・学術論文の被引用数

学術面でのアウトカム評価指標のひとつとして挙げられることが多い。しかしながら、全ての厚生労働科学研究について、系統的に把握するためには費用や労力が大きい、日本語論文について系統的に被引用数を把握することが困難である、研究領域・その領域の研究者数によって格差が大きい、自己引用や自らの雑誌の論文を引用するように推奨している雑誌もあり、恣意性が除去しきれない、学術論文の被引用数の重視が進むと、行政的な成果が軽視されるおそれがあるなどの課題がある。

・マスメディアへの掲載数

社会へのインパクトを定量化する指標のひとつとなりうる。米国AHRQにおけるプロジェクト評価指標として、マスメディアのデータ

ベースを解析して定量化する手法の検討が始まっているが、まだ実用化の目処は立っていない。また、メディアの関心度と、健康への影響の大小とは異なることが多い、問題発生を未然に防止した場合にはメディアに取り上げられることは少ないなど、研究の成果以外の要因によって左右されやすいなどの課題がある。

2. アウトカム評価の方法

(1) 基本的な考え方

アウトカム評価は、研究成果がその利用者どのように貢献したか、学術的、行政的にどのようなインパクトを及ぼしたか、国民の健康・安全を向上させたか、といった側面から研究活動を評価することを目的としていると考えられ、その目的に整合する方法とする必要がある。

(2) アウトカム評価として考えられる方法

一般論として、研究者からの報告、行政担当者からの報告、ピアレビュー、データベースの検索、統計資料の解析などの方法が考えられる。また、全ての研究を対象に毎年制度的に実施する方法と、臨時的にアウトカム評価を実施する方法などがありうる。

(3) 米国における状況

米国におけるFaculty workload survey (2012年) によって、研究者の勤務時間の42%が研究以外の業務に忙殺されている状況が明らかになっており、研究者の事務負担の軽減が提唱されている。米国NIHやAHRQでは、文献・マスメディア等のデータベースを活用したフォローアップ評価が実施・模索されており、研究終了後にも研究者からの報告を求めたり、ピアレビューを行ったりということはされていない。研究者や行政担当者に過度な負担をかけることは、研究活動の抑制につながり兼ねないため、留意が必要である。

(4) 推奨される方法

制度形成への貢献などについては、引き続き研究者から、また必要に応じて行政担当者からの報告を求める意義は大きいと考えられる。一方で、諸外国や多分野における実施状況を注視しながら、過度な負担にならないよう、データベースの検索や統計資料の解析などを中心にアウトカム評価を実施することが推奨される。また、評価指標については、可能な限り、客観的評価ができる(数値化された)指標を用いることが望ましい。

3. 良いアウトカムが得られる研究

(1) 目的、ゴールの明確な設定

良い研究成果が出されるためには、まず公募の際などにおいて、求められる研究成果を定性的・定量的に明示することや、研究成果が行政施策や国民の健康・安全・福祉の向上にどのようにつながるかを明示することが必要である。

(2) 行政的な成果の推進

厚生労働科学研究は、一義的には行政的な活用を目的とすることが多いため、行政的な成果をあげることは重要である。自殺対策のための戦略研究においては、研究者としての学術的な目標と、行政担当者の考える目標、また外部評価者が重視する市民への成果報告の目標とが、それぞれ異なっていた。このようなことは、他の研究事業においてもよくみられる状況であると考えられる。それぞれの立場が異なることから、当初考えている目標は異なって当然であると考えられる。しかしながら、研究者と行政担当者のコミュニケーションを密に行うことによって、異なる立場の者が重視する目標を理解し、収めんした目標を形成することが重要であると考えられる。具体的には、行政担当者が研究班会議に出席し、研究の進捗状況を把握するなど、研究者と顔を合わせてコミュニケーションをとるこ

とが有用であろう。また、研究途中で行政担当者が人事異動となった場合に、早期に行政担当者と研究者と連絡を取り合い、研究の趣旨や進捗状況などについて確認することが有用であろう。

(3) 学術的な成果の推進

行政的な目標のための取り組みについても、諸外国や国内の各地域において同様の課題を抱えていることも多い。また、時代を超えて、その根拠を確認する必要が生じることもある。その意味では、一義的には行政的な成果のための研究であっても、学術論文や学会発表等の学術的な成果として国内外に発信し、また後生に残すことの重要性は大きいと考えられる。

(4) 社会への情報発信の推進

厚生労働科学研究は貴重な税金によってまかなわれていることから、研究成果について、より一層、社会に情報発信を行っていく必要がある。数量的な評価結果だけではなく、社会に貢献した事例や、ある症例における研究の恩恵など、具体的で一般国民にわかりやすい形での説明も重要であろう。また、特に、厚生労働科学研究は、制度形成を通じて国民の健康に貢献する研究が多いことから、そのような貢献についても情報発信していくことが求められる。大学等においては、教員の任用や昇進において論文数や被引用数などの学術的な業績を重視しているところもある。一方で、行政的な成果によって国民の健康に貢献することも重要であることから、研究推進事業の活用なども行いながら、それらの業績やその重要性についても積極的に情報発信する

必要がある。

(5) 体制の構築

行政的な成果、学術的な成果、社会への情報発信について、時間や労力的、また能力や適性として一人で全てをきちんと担うことは困難なことが多い。そこで、研究班を組織する際には、それぞれに秀でた複数人で協力してそれぞれの機能を分担することができるチームとして体制を構築することが必要である。また、研究の本体部分が終了した後に、学術的な成果や、社会への情報発信等を行うには、一定の時間が必要なことも多いため、今後、それらを推進する研究支援体制をより一層充実させることも必要であると考えられる。

E. 結論

厚生労働科学研究がより効果的に実施され、国民の健康等にさらに一層貢献できるように、公募の際などにおいて、求められる研究成果を定性的・定量的に明示して、それに沿ったアウトカム評価を推進していくことが必要である。

F. 研究発表

- | | |
|---------|------|
| 1. 論文発表 | 該当無し |
| 2. 学会発表 | 該当無し |

G. 知的財産権の出願・登録状況(予定を含む)

- | | |
|-----------|------|
| 1. 特許取得 | 該当無し |
| 2. 実用新案登録 | 該当無し |
| 3. その他 | 該当無し |

参考資料 1 米国におけるアウトカム評価に関するインタビュー記録

実施日 2016年2月2日(火)

場所 Center for Scientific Review (CSR), National Institute of Health (NIH)

インタビュー対象者:

Richard K. Nakamura, Director, CSR, NIH

Donald Lockett, Communications Director, CSR, NIH

Mike Lauer, OER, NIH

Robert M. Kaplan, Chief Science Officer, AHRQ

NIH: National Institute of Health (米国国立衛生研究所)

CSR: Center for Scientific Review (科学審査センター)

OER: Office of Extramural Research (外部委託研究局)

AHRQ: Agency for Health Care Research and Quality (保健医療向上研究機構)

NIH 及び AHRQ とも、DHHS: Department of Health and Human Services (保健福祉省)内の組織

Part 1. NIH 科学審査センター責任者 Richard Nakamura 氏からの概要説明

専門家による評価、検証された研究の価値(パワーポイント資料参照)

1. NIH は、2500 研究機関、30 万人に及ぶ科学者並びに研究に携わる職員を支援
2. NIH 科学審査センター
 - i. NIH 科学審査センターが、全ての NIH 助成金の申請を受理する。
 - ii. 6 万の申請の内、75 パーセントを検討する。
 - iii. 一年に 1 万 7 千人の審査委員を募集する。
 - iv. 一年で 1,500 回、審査委員会が開催される。
 - v. その管理運営の為、247 人の科学審査職員を要している。
3. NIH 助成金申請の専門家
4. 審査方法並びに資金源
 - a. 国立衛生研究所(NIH)
 - b. 科学審査センターが、NIH 諸機関並びに専門家審査グループに対し、仕事の分担を決める。
 - c. 提出された申請書を科学的に検討
 - d. 研究の優先順位に照らして評価
 - e. 諮問機関の推薦
 - f. 総責任者が最終的決定を下す。
5. 評価方法並びに責任の所在
 - a. 手順一 SRO (Scientific Review Officer) を通して NIH

- b. 科学的一議長及び研究分野の専門家委員
 - c. 専門家による評価の重要性
 - d. 助成金を出す決定は NIH 諸機関
6. 科学的評価委員の役割 (評価手順全体の責任を担う連邦政府に任命された専門家)
- a. 申請書類に目を通して、書類に不備がないかを確認
 - b. 評議会の意見も参考に、委員長並びに評価委員を選任
 - c. 議長並びに評価委員の CSR 教育訓練の実施
 - d. 評価後の要約書類の準備
7. 評価委員長の役割 (経験豊富な学外の科学者が委員長になり、科学審査委員達と協力して会議を進める。)
- a. 委員会の進行に務め、かつ、委員同士の話し合いを要約する。
 - b. 全ての委員の意見を十分尊重する。
 - c. 会合において、科学的意見交換を促進する。即ち、予定表に則り、話し合いが取り落としのないように努める。
 - d. その科学分野における代表者としての顔
8. プログラム職員の役割
- a. NIH の各研究所 (Institutes and Centers, IC) やプログラムに割り振られた申請をモニターする。
 - b. 評議会や各研究所に受賞の優先順位を推薦する。
 - c. 各研究所のミッションを推し進める為に、受賞者の一覧表を作成
 - d. 申請者の問い合わせや相談に積極的に応じる。
9. 評価委員
- a. 関係する研究分野における博士号或いは医学博士取得者
 - b. 独立した研究並びに競合する研究分野で成功歴を保持
 - c. 優秀な文献の公表
 - d. 科学的分野での承認及び専門家として認められている。
 - e. 真摯で公平な評価
 - f. 評価における独立性
 - g. 多種多様な見方
10. 評価委員の評価を超えて
- a. プログラム職員を通して諸機関が、機関のミッションに照らして検討後決定
 - b. 助成金管理運営職員が実際に助成金を受渡し、条件付きの場合はその条件を明示する。
 - c. 研究代表者の機関が、研究費並びに間接費を含めた助成金を受理し、そしてその双方を管理運営する。
 - d. 研究代表者側が、進捗状況をプログラム職員に報告する。
 - e. 研究代表者は研究結果を公表する。
 - f. 研究代表者は助成金の良し悪しを検討の上、最終報告書にまとめる。
11. なぜ、NIH 専門家評価手法が成功しているか。
- a. 申請者に明瞭透明である。
 - b. 機関そのものではなく、むしろ、助成金の行く先を、概念や個人に焦点を当てている。
 - c. 米国全土の独立した研究者からアイデアが生まれでる。

- d. 起業家同様、資金獲得の為、研究者は競争することが重要である。
 - e. 主な評価委員は外部地域の科学者がなる。
 - f. 科学者並びに係の職員は、公正を期するため、最善の努力をばらう。
12. NIH の評価手法を経た研究が利益及び効果を生み出した理由
- a. NIH 研究の影響 ― 詳細は www.nih.gov/impact を参照。
 - b. 成果の歴史 ― 148名に及ぶ、NIH が支援しノーベル賞受賞者数
 - i. 2015年の受賞者 3名 ― 科学分野、Aziz Sancar と Paul Modrich、―― 経済学、Angus Deaton
 - c. 経済的利得 ―
 - i. 2012年、全米で、NIH の学外助成金が570億8千万ドルの経済生産高を上げた。
 - ii. 年平均 3 兆 2 百億ドル
 - 1. 研究成果の指標の一つとして、1970 年―2000 年の間に、寿命が伸び、経済的価値に換算すると、約95兆ドルにおよぶ。
 - d. 公共の研究開発への投資と民間の特許取得
 - i. 研究分野に対する NIH の資金提供で、その分野の民間特許が、2.3 追加される。言い方を変えると、2 つか 3 つの NIH 助成金毎に、1 パーセント民間特許取得が増える。
13. 科学の進歩及び健康増進
- a. 生物医学の進歩及び社会健康増進 ― 米国人の寿命が延伸
14. NIH の研究は多大な成果を上げている。
- a. NIH の研究成果が、心臓血管の病気、脳梗塞、癌、糖尿病等の治療法の躍進に貢献、毎年135 万人の命を救うことに役立っている。
 - b. 米国議会報告書、「医学研究並びに NIH の役割」の2000年版によると、NIH が資金提供した研究により、主な製薬の70パーセントが開発された。
15. NIH の専門家評価に関する新しい統計
- a. 研究成果の引用数と専門家評価回数とは深い相関関係がある。
16. NIH の専門家評価手法を評価することも重要

Part 2. NIH 外部委託研究局 Mike Lauer 氏との質疑

専門家の評価、

専門委員会の評価点と助成金授与の結果との関連を調べた。特に、公表回数、引用回数等と言った主な指標に注目した。もう一つは、臨床試験の成り行きに注目、いかに早くその主な結果が公表されたかを調べた。臨床試験の成果が多量である場合、早く公表される場合が多い。

専門家委員会の評価が、普通程度或いはそれ以下の場合、引用数も少ないといった相関関係があることが判明した。要は、評価点と引用数には緊密な相関関係が見られる。

もう一つの問題点は、成果の成功率のみを考慮するのではなく、金額支出を考慮することも重要だ。興味深いパターンが見えてきた。金額支出高が多いほど、専門家委員会の評価も高い傾向がつよい。例えば、X 軸に予算を、Y 軸に専門家委員会評価点を取ると、予算が多い程、評価点も高いと言う、両者の緊密な関係が明瞭に現れる。

その結果、専門家委員会評価点が高い研究分野に、助成金を回す傾向がより強くなる。それはそれで良いが、高額研究分野に助成金がより多く流れると、資金提供できる研究分野の数がより狭まる。

その上、研究成果のみならず、出費した額に応じて、どの程度の成果があったかを考慮することも重要だ。資金を提供する側としては、科学そのものの質の高さばかりではなく、その成果を出すために、どの位の出費が必要であったかを熟慮することも大切だ。

引用数と公表率は費用対効果の指標の一つではあるが、国立癌センター及び CDC の Muin Khoury 氏が研究成果測定指標として、PQRST をまとめた。

P (Productivity)、生産性、すなわち、公表数、引用数等

Q(Quality)、品質、研究内容の質の高さ

R(Reproducibility)、研究成果再現の可能性—これは、最近、NIH では大変重要な問題点になっている。

S(Sharing)、研究結果が他の科学界でも参考になる度合い—これも、現在、NIH の重要点だ。

T(Translation)、研究の結果が実際の患者にどのように役立つ形で実現化しているか、また、医療政策に反映したかも重要だ。

研究分野によっては、このような指標でわりと良い結果を導きだせる。

臨床試験では、研究成果の良さを評価する客観的項目があるが、最終的価値判断は、誰かが主観的価値判断をする必要もでてくる。例えば、臨床試験の手順ややり方の適応性は誰かが判断することになる。

米国政府の一部である AHRQ (Agency for HealthCare Research and Quality) では、詳細に科学的調査研究が行われている。そこでは、誰が資金提供をしているかは無視して、研究そのものの質の優劣に関して、組織的に格付けが行われる。これも一種の測定手法である。

我々や他者が資金提供した研究分野の成果を再現できない、或いは、再現し難いと言う問題点が持ち上がっている。その理由は、研究手法のデザインに問題点があった場合もあるし、研究者が不十分な情報しか出していない時もある。

そのため、他の研究者がその研究結果を再現しようとしても、困難である。であるから、その問題に対処するため、NIHでは助成金の申請書提出の際、研究のデザイン、再現のしやすさに関連する項目を追加した。

2016年1月に始まったばかりであるが、初期段階の申請書提出の際、再現可能性に関する要項書き込み欄も追加した。そして、NIHのプログラム系の職員、評価委員にこの件に関する情報を周知、教育訓練強化を図ることになった。

データであれ、細胞系、ネズミ系であれ、最終的に我々が求めるのは、研究成果を分け合う可能性の重要さだ。他の科学者がどれくらいその物質を必要とし求めるかも重要だ。研究成果の情報を提供する意思があれど、誰も興味を示さない場合は問題だ。

反対に、データをシェアする意思があり、しかも、大勢の研究者が関心を示し、そのデータを使って、自分の研究にと積極的に取り入れる場合は良好だ。

長い間、NIHではデータシェアに関心があり、例えば、NIH 心臓血管研究所においては、臨床試験のデータシェアが、過去20年以上長きに渡り実施されていた。

その上、7、8年前から、遺伝子関連分野に関するデータシェア政策を新たに作り、それは、実際とても上手く稼働している。数千人の科学研究者が遺伝子データをNIH研究者と分け合っている。それを元に、数千の論文、文献が発表されている。

いま創意工夫中の事項は、他の研究分野においても、データシェアの可能性もあるか模索中である。遺伝子関連分野に限らず、基礎研究成果のデータ、臨床試験結果のデータ等のシェアが可能であるかも調べているのだ。

研究成果の測定手法の一つとして、データシェアの高さも考慮するということだ。

研究成果の発表、特許取得件数、また、科学界でどれほどデータシェアがなされたかも全て重要項目なのだ。

翻訳、要約、政策への反映度だけで判断すると、基礎研究分野には不利に働くことも重々考慮する必要がある。一方、第3相の臨床試験等では、大変公平なやり方である。試験の結果、治療法の改善、医療政策の見直等につながる可能性が高い。

このように、多種多様な方面から判断を下すことが重要だ。

勿論、色々評価手法があるにせよ、専門評価委員の良好な審査が重要であることを忘れてはならない。まだまだ、このような審査や評価手法は、完全に科学的とはいえず、常に改善策を模索中である。

他の評価法

我々は、SPIRESと言う、公表された論文、研究成果のデータを保持している。この中に、170万にも及ぶ論文がある。これはNIHが支持しているものだ。これは、1980年の中頃から収集された。本当はこの数字よりもっと、多大だと思う。私の予測はあくまで低めに抑えている。

特に、10年、15年以上前は、論文数も本来より少なめにしか収集されなかった可能性もあるからだ。

また、NIHは引用に関するデータも集めている。

最近、新しいやり方、Relative Citation Ratio(相対的引用率)も開発された。これは、無料でだれでも得ることができる。これで、同時期に、類似の論文と比較して、論文がいかに多く引用されたかを測定することができる。

また、我々は、下記のような評価プログラムも取り入れている。具体的例を一つ上げると、最近、訓練プログラムを実施した。若手の研究者に助成金を授与するプログラムで、その為の教育訓練がおこなわれた。

その時、訓練に参加した若手研究者がその後、実際助成金授与に成功したかを確認した。論文、文献の発表が実際あったか、回数が増えたか、科学分野で積極的に活躍し続けたか等も調査した。

また、数年前、Pioneer Awardsプログラムを実施した。特に優れた科学者に対して授与するグラントで、プロジェクトそのものの良さ以上に、科学者自身の優秀さ、生産性の高さを計る。我々が実際に評価を行ったプログラムだ。

また、NIH評価の一部として、National Bureau of Economic Research (NBER)、国立経済性研究調査局がNIHをしばしば調査している。今年の始めに、報告書を公表した。NIHの資金獲得レベルと特許取得件数の間に、良い相関関係があることを報告している。NIHプロジェクトが特許取得に大いに寄与しているのだ。

また、NIHグラントとHoward Hughes グラント、これはプロジェクトに対してではなく人に対して授与する賞であるが、この二つを、国立経済性調査局が比較検討した報告も出ている。

部分的には、ハワードヒュー賞の方が結果がより良い場合もあった。その点を改善する為に、NIHは科学者自身により重きを置いた授与を試みている。

国立衛生研究所内の各部署で、独自の評価法を取り入れている場合もある。その上、評価を専任とする職員を配置している。

最近、この分野に注目が集まってきている。NIHの戦略計画においても、数ヶ月前に、米国議会に意見書を提出した。その内容は、今後、「NIHはより結果に重きを置いて評価する」という主旨だ。評価法をもっと組織的、合理的に行うという意見書だ。

長期的追跡調査はNIHの場合していないが、最近、色々な角度から科学者の経験に対して、次のようなアンケート調査を実施した。例えば、グラント授与までの全ての行程に関する質問も含まれている。

FDP(Federal Demonstration Partnership)という連邦政府の部局があり、そこでは、詳細な質問調査が行われている。例えば、科学者に、連邦政府と取り組む必要ができた場合、どんな問題点があるかといった質問が含まれている。それも、大変詳細な調査である。

我々も、FDPと関わりが深く、FDPと協力している。言い方を変えると、我々側が、FDPに情報提供をするということだ。もちろん、予算はそれぞれ別々であるが、協力関係にある。

我々の注目している関心事の一つは、科学者の追跡調査を実施することだ。特に、懸念していることは、若手科学研究者が、その後、いつの間にか影もなく消えてしまう場合だ。再び、戻って来ない場合がある。二度とグラント申請もしないし、当然、授与されることもなく姿を消してしまった。

そのような科学者の実情を把握する必要もある。研究自体に興味をなくしたのか、他の職業に鞍替えした為、我々の元から消えたのか等、原因を究明する必要もある。

今、米国で大きな問題になっている点は、科学研究者に対する書類作成等といった事務的仕事量が増えすぎていることだ。先ほど話したFDP局の調査結果で、科学研究者は40%以上の時間を、事務処理に費やしているとある。60%のみ専門分野の科学研究に集中できるのは大きな問題で、今後この問題点の解決策に取り組む必要がある。

米国上院議会の要望で、全米科学財団が、小委員会を設置した。この小委員会は、全ての米国連邦政府に対して、幾つかの意見書をまとめ提出した。書類作成手順の単純化、合理化、書類の統一性の必要性等を推奨している。

注. Federal Demonstration Partnership (FDP) 2012 Faculty Workload Survey Research Report

http://sites.nationalacademies.org/cs/groups/pgasite/documents/webpage/pga_087667.pdf

正確には、研究以外に費やしている業務時間は42%であり、研究は58%という調査結果となっている

政府から特別の許可がない限り、我々は人々に対してアンケート調査をすることが許されていない。その理由は、研究者に何の利点もないのに、政府があまりにも小うるさくつきまとうという印象を忌避したいためだ。

アンケート調査をしようと思えば出来なくはないが、付加的な事務処理が必要になるのだ。

ORCID (Original Researcher Citation Identification Directory)、全ての科学研究者には背番号が付与されている。その科学研究者に何事が起ころうとも、例えば、他の大学に移転しても、民間企業に就職しても、他国に行っても、他の機関から資金提供を受けても、どの科学者であるかを簡単に追跡できる。論文やグラントからも追跡調査ができる。今週、実際にロンドンで、ORCID の会議が開催中だ(2月第1週)。

NIH に、研究者がグラント申請の為、論文発表等バイオ スケッチ(研究分野の略歴)を含めた研究課題提案書を提出する。

新しいプログラムである、R35 はグラントの一種で、以前の研究成果をもとに、科学研究者が自選でグラント申請することができる。研究者が今何に関心が深く、将来どのような研究をしたいと考えているかも考慮される。でも、大部分は科学研究者自身の質の高さに重きを置いている。

このプログラムはまだ始まったばかりだ。どのような結果になるかは、今後の結果を待つ必要がある。

長寿と科学研究の関連性に関して、幾つかの興味深い文献が出ている。数年前に発表された文献で、賛否両論を巻き起こした論文であるが、Manton 氏が主な執筆者だ。NIH において、違った病気ごとに、資金提供受率と病人に対する治療法改善成果との関連性を調査した。

病気の種類によっては、はっきりと関連性が出ることもある。例えば、心臓病の場合、NIH が心臓病に資金をつぎ込めば、つぎ込むほど、心臓病改善に繋がるといふ相関関係がはっきり現れる。他の病気では、それほど相関関係がはっきりしない場合もある。例えば、癌の場合、はっきりした相関関係は出にくい。

もう一つ関心の深い研究結果がある。それは、ハーバード大学労働経済学部の David Cutler 教授の論文だ。カトラー教授が主幹で、他の研究者と一緒に、40年から50年にも渡る長期的見地から、心臓血管病と治療方法による、結果の違い等を詳細に調査した。例えば、喫煙は心臓に悪影響を及ぼし、心臓病になる確率が上がる。その為、禁煙に努める。また、高血圧は病気の原因であるから、血圧を下げる治療が行われ、その結果、心臓病が減る。

カトラー教授らは、資金調達と病気削減の関わりを詳細に調査した。医学研究成果があまり病気を減らさず、あるいは、患者の改善にほとんど効果がないということを証明することに重きを置いた研究だった。

それにもかかわらず、仮に、研究費への投資が1ドルとすると、30ドル分の医療改善に繋がったという結論がでた。

心臓病改善の要素は多々あるが、コレステロール値を下げる、タバコを吸わない、適度の運動、血圧を下げる等や、アスピリンを飲んだり、血液凝固を避ける薬を飲んだり、手術を含めた色々な治療法を試みることで、心臓病改善の可能性もある。このような全ての要素を調査研究対象にしたのだ。

Part 3. NIH Richard Nakamura 氏、Mike Lauer 氏、Donald Luckett 氏との質疑

NIH では、品質改善の為、色々な調査研究を行っている。科学の政治的影響に関して日本側から質問を受けているが、一つ私が述べたいことは、政治的干渉を最小限に抑えるため、専門家評価は二段階を踏む手法を使ってグラント授与を決定している。

経験豊富な科学者が評価をすることで、評価そのものに米国議会が関心を寄せても、評価法が実際は大変複雑で、議員の手に負えないことに気づく。結果的に、NIH の上層部で管理している科学者の意見を尊重することになる。

70年にも及ぶ長期間の科学成果を元に、科学の基礎研究の重要性が深く認識されている。

「応用科学や臨床試験にもっと焦点を当てよ」と言う、意見もない訳ではないが、基礎研究の成果が、多くの科学研究向上に貢献している事実も見逃せない。病気の改善、治癒や、コンピューターに至るまで、幅広い影響力を持つのが科学基礎研究なのだ。最終的に、米国の強い味方になっている。

基礎研究の重要性を示唆する事例として、例えば、日本も、基礎研究に多くの成果を上げている。でも、我々側は、全員一致協力して、基礎研究の重要性を示唆し続けることが重要だ。

民間企業は、基礎研究にあまり積極的ではない。政府のみが、基礎研究の重要性を認識し、後押しをする。過去の業績、成果で理論武装して、例えば、1ドルの投資が50ドルの見返りになるといった説得も必要だ。勿論、計算法は単純ではない。

一般的に、米国議会は我らの説明に賛同している。時々、ある議員が反論することもあるが、NIH は全体を通し説得力を有している。

もう一つの大きな利点は、NIH の決定権が分散していることだ。一人の人が、全ての決定権を掌握していない。最良でない決定がなされても、その責任の所在を特定し難い。NIH の組織の中で、一人を解雇しても、NIH 全体に影響が出ることはない。

であるから、科学に対する資金提供は、多くの人々が、色々な決定をして確定する方式になっている。一人の人間が、組織に壊滅的打撃を与えることは不可能だ。これが政治的に上手くいった理由だ。

政府そのものは、上下関係がはっきりしている組織であるが、NIH はもっと水平な組織になっている。

あるレベルの評価がある。一般的に、助成金を授与した研究は、当然、例えば、論文発表といった、何らかの結果を生み出す。研究分野が狭すぎる事例もある。契約した場合もある。でも、それはまれである。例えば、一万人の遺伝子情報を収集する。そのような場合でも、科学的に最良な、「一万人を選抜する方法は如何に」と、いった観点から見る場合もある。他の研究者のやり方と比較することで、競争できる分野だともいえる。

「ある特定の人しかできない」ということはまれだ。全て競争することが可能だ。最終的評価は、最初の目的を果たすことができたかで判断できる。

30億ドルから40億ドル位までは、契約で研究がなされる。科学調査者が積極的に案やアイデアを自分から提出する方法を採用している。NIHが、例えば、一つの病気に限定した狭い研究、具体例はアルツハイマー病を挙げると、広く研究を奨励している分野だ。アルツハイマー病関連ならば、どの内容でも研究対象になり得るのが現状だ。色々なアイデアを提案することができる。

一番良いと信じる研究対象、アイデア、どの様なやり方を採用するかといったことも、学外の研究者次第である。評価委員達を、如何に説得するかといった科学研究者の説得力が重要だ。

そして、その結果が上手くいったかを調べる。重要な点は、危険をものともせず、チャレンジ精神を持って、新しい研究に取り組む姿勢が重要だ。研究プロジェクト全てが、常に成功すると仮定すべきではない。良い結果を期待するには、ある程度危険を冒す覚悟も必要だ。

上層部に、危険を承知で、新しい研究に取り組む必要性を説得することも重要なのだ。選択肢はしばしば複雑である。

日本も同様であるが、NIHでは、何十年もの長きに渡り、継続的に助成金提供を継続してきた歴史があり、投資の見返りが十分あるという証拠がある。また、助成金付与で、科学界の活性化に寄与してきた。また、良い科学を選別する手法も経験で学び取ってきた。科学社会の清浄化にも一役かっている。生産性の低い科学者は排除、生産性の高い科学者の後押しに力を注いだ。

政治の世界でも、理解が深まっている。例えば、自動車業界は、世界中を市場としているため、生産拠点を分散する傾向が強い。NIHの場合は、研究助成金を全国に授与している。投票権を有する市民を通して議員に影響を及ぼし、NIHの重要さを議員に認識して貰うためにも、広く全州に助成金を授与することが重要だ。

議員は、自分の選挙区内にある大学機関が授与することで、議員のNIHに対する認識も深まるのだ。

責任の所在に関して

ある論文を上げて、この研究結果が資金提供の見返が多いと指摘することは可能だ。でも、研究をすればするほど、治療法の種類も増え、医療費高騰に繋がると指摘する人々もいる。結果的に、国にとってまた国民にとっても、費用がかりすぎるという議論だ。それにも一理はある。