

奨励金No.1523

若手医学研究者の立場から、より良い医療を目指すために 「医学研究・研究者の評価軸」を再考する

杉山 雄大

国立国際医療研究センター研究所糖尿病情報センター 医療政策研究室長

Reconsideration of evaluation criteria for medical research and researchers aiming for better medical care from the standpoint of young medical researchers

Takehiro Sugiyama,

Diabetes and Metabolism Information Center, Research Institute, National Center for Global Health and Medicine, Chief, Division of Health Services Research



本研究は、日本の医学研究コミュニティにおける研究評価に対する認識の一般的傾向と、コミュニティ内部に存在する差異を明らかにすることを目的とした。多様なキャリアや専門の研究者に対する聞き取り調査、医学系各学会を通じたアンケート調査が実施された。研究結果から、医学研究評価においてインパクトファクターなどの定量的指標を重視する一般的な傾向が確認された。研究分野や世代、性別などに応じて研究評価において重視する要素が異なることも判明した。

The purpose of this study was to identify general trends in perceptions of research evaluation in the Japanese medical research community and the differences between the sub-communities. We conducted interviews with researchers from diverse careers and specialties and distributed questionnaires through medical societies. The study results confirmed a general tendency to emphasize quantitative indicators such as journal impact factors in medical research evaluation. It was also found that the factors emphasized in research evaluation varied depending on the research field, generation, gender, and other factors.

1. 研究内容

1.1 研究の位置づけ

研究を評価することは、国内および国際的に科学の進歩を継続的に進めるために不可欠である。これまで英文論文数や被引用数などの定量的指標は、資金配分や人事において重要視されてきた。一方で、このような研究評価の現状に対する研究者の意見については体系的な調査はほとんどなされてこなかった。

そこで本研究は、日本の医学系研究者の研究評価に関する認識を明らかにし、その問題点を抽出

することを目的とした。多様なキャリアや専門の研究者に対する聞き取り調査や、基礎医学、臨床医学、社会医学の各分野の医学研究者を対象にアンケート調査を実施し、現在の医学研究者の評価軸の特徴や問題点を検討し、今後検討可能な評価方法を明らかにした。

1.2 調査方法

図1に調査の概要を示す。本調査は、第31回日本医学会総会のU40委員会若手研究者有志が中心となり実施した。日本医学会分科会に所属する

研究者、計22人のシニアおよび初期キャリア研究者や学生に対して予備質問表の作成およびフォーカスグループインタビューを実施した。併せて、質問紙の検討のため、10名以上の研究評価に関する非医学系有識者から意見を聴取した。

完成した質問票を用いて、医学研究者を対象にインターネット調査を実施した。本調査は、日本医学会を通じて各学会に配布を依頼した。調査期間は、2022年12月14日から2023年1月17日までとした。

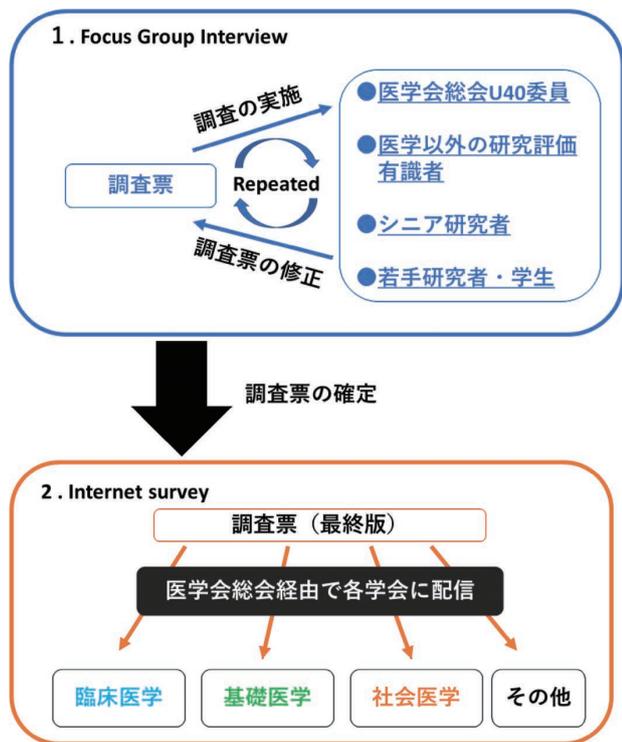


図1. 調査の全体像

1.3 結果

調査期間中、合計3169名の回答を得た。このうち、無効回答30名を除外し、3139名（男性2244名、女性852名、その他43名）の研究者を分析対象とした（図1）。

研究者を評価する定量的指標（図2）について、雑誌のインパクトファクター（IF）が重要視されており（特に重要：19.6%、重要：47.8%）、特に基礎医学では臨床・社会医学よりも、若手ではシニアよりも、男性では女性よりも高かった。英文誌に掲載された論文数（特に重要：33.3%、重要：51.8%）は、和文誌に掲載された論文数（特に重要：3.4%、重要：30.2%）よりも重要視されていた。日本語の雑誌よりも英語の雑誌を好む傾向は、臨床・社会医学分野よりも基礎医学分野で、60歳以上の研究者よりも39歳以下の若い研

問3-1.
以下の定量的な指標は、ご自身が周囲の研究者を評価する（人事に限らず、優れた研究者と考えるなども含む）際に、どの程度重視しますか。

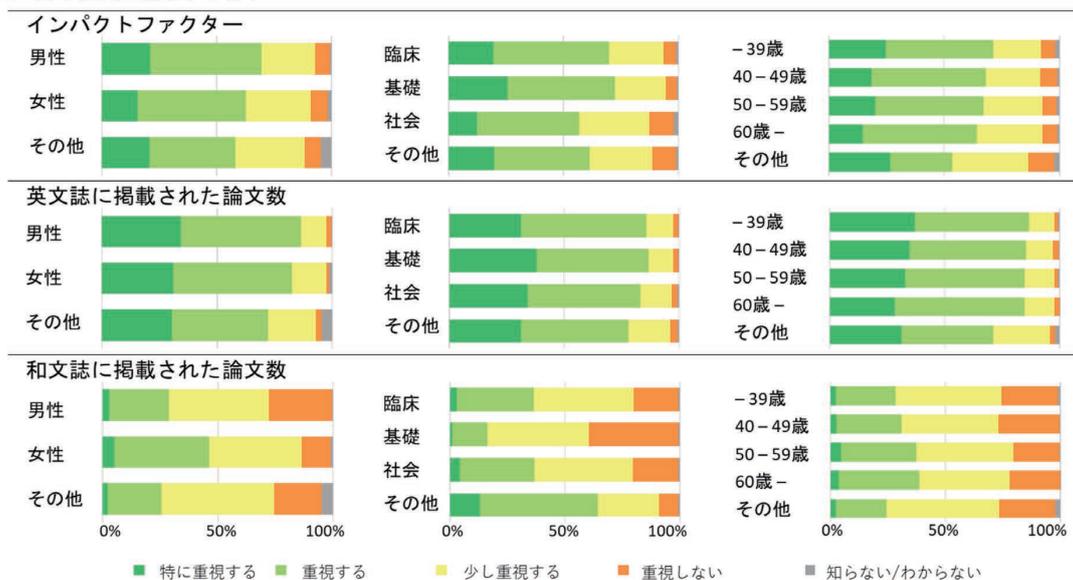


図2. 研究者評価における定量的指標に対する認識

究者で、女性よりも男性で顕著に見られた。

研究者を評価する際の定性的因子（図3）については、分析の網羅性（特に重要：12.6%、重要：51.6%）よりも、研究テーマの独創性（特に重要：37.0%、重要：51.5%）や科学の進歩への貢献（特に重要：37.4%、重要：47.4%）の方が重要だと考えられていた。研究テーマの独創性は、基礎医学では臨床医学・社会医学よりも重要視され（特に重要-基礎：46.2%、社会：39.1%、臨床：32.0%）、臨床・社会問題の解決への貢献は社会医学では基礎・臨床医学よりも重要視されていた（特に重要-基礎：30.4%、臨床：41.0%、社会：52.0%）。

図4は、研究評価のための定量的指標と定性的因子に対する研究者の認識を示したものである。特に40-49歳、男性では、基礎医学、社会医学において、定量的指標では適切に評価されない重要な研究があると考えられる傾向がみられた。定量的指標の使用は、臨床・社会医学の研究者に比べ、基礎医学の研究者にとっては不正行為につながる可能性があることと認識されていた。シニアに比べ若手研究者において、定性的評価は無意識のバイアス

に影響されやすいと考える傾向があった。

その他の要因で層別化した結果については、医療従事者以外に比べ、医師などの医療従事者は、IFが重要であると回答する頻度が高かった。雇用形態を考慮した場合、上級管理職は、定量的な指標は研究以外の活動を軽視していないと考え、対人関係や無意識によるバイアスに気づかない傾向があった。研究評価に関するサンフランシスコ宣言（DORA）を知っている研究者は、IFの重要性をやや低く評価する傾向があった。中でも、DORAの内容まで知っている研究者は、テーマや方法論の独創性、科学の進歩への貢献、臨床や社会の問題解決への貢献といった定性的な因子を重視する傾向があった。さらにDORAを知っている研究者は、1) 重要な研究でも適切に評価されないものがある、2) 定量的な指標は研究不正につながる可能性がある、3) 定性的な評価者（=査読者など）の評価の妥当性を評価すべきである、と回答する傾向にあった。

自由形式の質問に対しては、合計645件の回答を得た。これらの回答を、仕事以外の活動（n=11）、仕事における研究以外の活動（n=84）、指

問3-2.
以下の定性的な因子は、ご自身が周囲の研究者を評価する（人事に限らず、優れた研究者と考えるなども含む）際に、どの程度重視しますか。

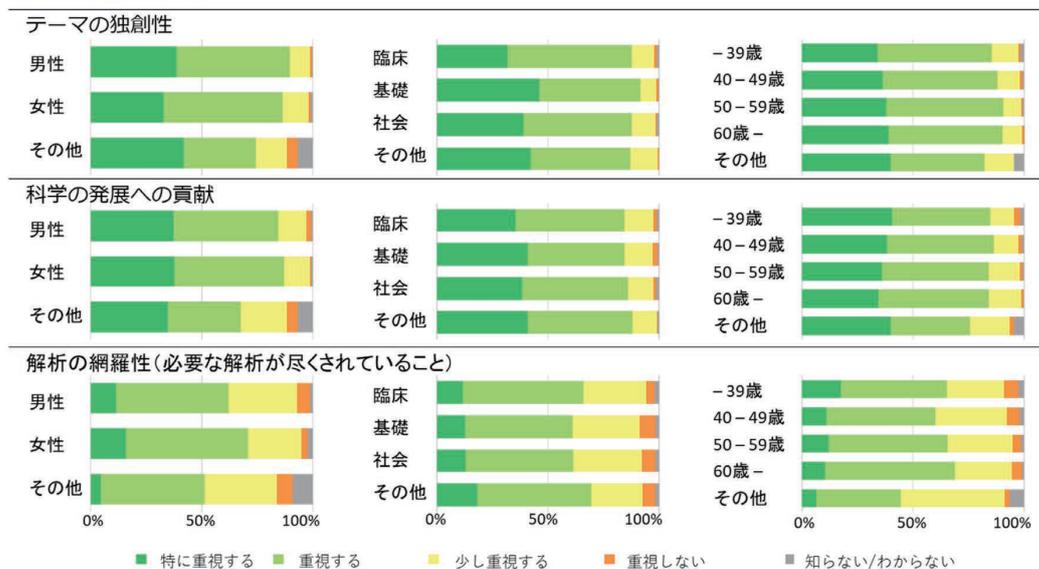


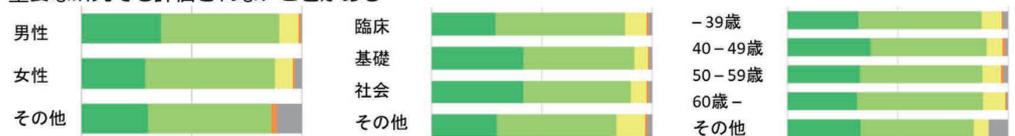
図3. 研究者評価における定性的因子に対する認識

問3-4.

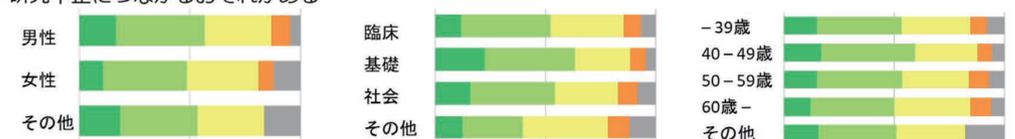
私たちの先行調査では、被引用数やインパクトファクターなどの定量的指標による研究評価について、様々な意見があることが分かりました。

定量的指標を評価に用いることに関する以下の意見について、ご自身の考えを教えてください。

重要な研究でも評価されないことがある



研究不正につながるおそれがある



問3-5.

私たちの先行調査では、広い意味でのピアレビュー（論文の査読に限らず、同じ専門分野の研究者が研究論文を読んだり学会発表を聞いたりして研究内容を評価する作業）などの定性的な研究評価について、様々な意見があることが分かりました。

定性的な研究評価を行うことに関する以下の意見について、ご自身の考えを教えてください。

アンコンシャス・バイアス（自分自身気づいていないものの見方や捉え方の歪みや偏り）の影響を受けやすい

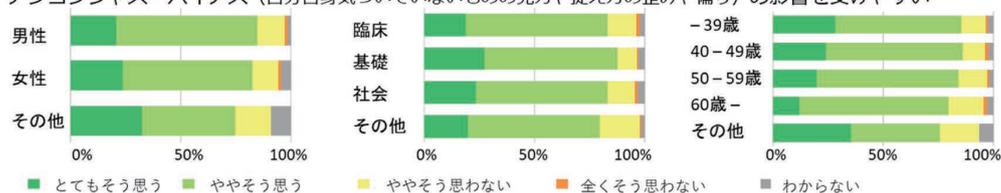


図4. 研究評価における定量的評価と定性的評価に対する認識

標の性質（n=69）、評価疲れ（n=6）、研究評価に関する制度・環境条件（n=77）、価値ある研究と評価の間の構造的葛藤（n=11）、本調査そのものに関する回答（n=232）の7つに分類した。

1.4 考察

日本の医学研究者、特に基礎医学、若手、男性の研究者の多くが、研究者を評価する際の定量的指標としてIFが適切であると考える傾向にあることが分かった。医学研究および研究者を評価する際の定量的指標および定性的因子に関する医学研究者の認識は、研究分野、年齢、性別といった特徴によって差があった。

医学系研究者におけるIFやその他の定量的指標を重視する一般的な傾向は、DORAやライデン宣言が懸念する状況であると言える。若手ほどIFを参照する傾向が強いが、それは多くの若手研究者が激しい競争下にあり、広く使われている評価指標を内面化する傾向があるのが一因と考えられる。

このような論文という評価軸を重視する傾向は、研究者の質的評価要因の中で、「分析の網羅性」が「研究テーマの独創性」や「科学の進歩への貢献」よりも低く評価されている本研究の結果とも整合的である。すなわち、業績を論文というひとまとまりの形で提示し評価において重視される慣行が一因であると考えられる。

また、研究分野、年齢層、性別で評価軸が異なることも特徴的であった。例えば、基礎医学の研究者は、IFや英文誌に掲載された論文数を高く評価し、和文誌の論文数を低く評価する傾向がある。これは、基礎医学の研究成果はどの国でも通用することが多いため、英語で情報発信することが合理的であるためと考えられる。年齢による認識の差は、評価者と評価される側の対比を表している部分もあるだろう。シニアが若手研究者を評価することが多く、評価軸を決定する権限を持つことに注意が必要である。また、女性は男性よりもIFを重視しないなど、性別による軽度の違いが見ら

れた。研究評価について議論する際には、性別の多様性を考慮する必要があると考えられる。以上のように医学系研究者の内部における評価要素が多様であることを前提に、評価基準に対する不断の見直しが求められる。

研究の限界には、自記式アンケートを用いたことが挙げられる。また、今回のサンプルは3,000人以上の回答があり、ある程度頑健な結果が得られたが、包括的調査ではない。研究評価に興味がある人ほどアンケートに回答しやすいため、回答者の属性や回答内容に選択バイアスが生じた可能性がある。しかし、本研究は、日本全国の医学研究者が研究評価や研究者評価をどのように捉えているかを調べた初めての研究である。本研究の成果は、研究者の評価方法の改善、ひいては研究成果の向上につながることを期待される。

2. 発表（研究成果の発表）

箕浦明. 「U40 委員『将来に向けた医学研究評価のあり方についての調査』報告」、第31回日本医学会総会（東京、2023）

（本調査は、第31回日本医学会総会の上記シンポジウムで議論を行うことを目標として、日本医学会総会学術委員会、日本医学会、日本医学会連合等の協力、指導を得て行った。）

杉山雄大. 「医学研究・研究者の評価の在り方について考える～第31回日本医学会総会 U40 委員会アンケート調査から～」、第75回日本産科婦人科学会学術講演会（東京、2023）

箕浦明. 「『将来に向けた医学研究評価のあり方についての調査』の結果報告と今後の課題について」、日本医学会連合第7回社会医学若手フォーラム（オンライン、2023）