

安易な数値頼みの研究評価からの脱却を

Atsuko TSUJI **辻 篤子** 科学ジャーナリスト, 中部大学 特任教授



科学の信頼性をおびやかすところまできた

「科学者をまどわす魔法の数字, インパクト・ファクターの正体」(日本評論社)という本が今年初めに出版された。

インパクト・ファクター (IF) とはいうまでもなく、学術誌の重要度を示す1つの指標である。科学者ならおそらく知らない人はいないだろう。掲載された論文が多く引用されるほど注目度が高い、つまり、文字通りインパクトが大きいことを示す。「正体」とはおどろおどろしいが、IFの算出方法や研究分野の違いによる影響などを詳細に見た上で、副題にあるように「誤用の悪影響と賢い使い方を考える」のが本書の狙いだ。

須藤靖・東大教授(物理学)は朝日新聞の書評で、「IF中毒の哀れな科学者のみならず、物事を一次元的指標でラベル付けして思考停止に陥りがちな現代人の必読書である」と評した。問題の根は深い。

IFの偏重、とりわけ研究評価に用いることの問題点はかねて指摘され、警鐘も鳴らされてきた。しかし、歯止めはかからず、ついに「科学の信頼性をおびやかすところまできた」というのが著者である麻生一枝・成蹊大非常勤講師の感慨だ。動物行動生態学者として在米約20年、帰国後に大学の研究現場で感じた違和感が執筆の動機になったという。世界的には見直しの動きも出てきたという今、改めて考えてみたい。

研究分野によって顕著な差がある

学術誌のIFは、過去2年間に掲載された論文がその年に引用された回数の平均値と定義される。「Nature」「Science」「Cell」といった雑誌では40近くになり、超一流誌とされる一方、多くの学会誌は1前後だ。

そのIFが雑誌の評価にとどまらず、研究者の評価に直結し、人事などの際に重要な要素になることが少なくない。論文が掲載された雑誌のIFを合計した数字で研究者の能力を評価するといった乱暴な例も、さすがに減ってきたといわれるものの、決して珍しくはない。

なぜ乱暴かといえば、まず、IFはあくまでも平均値であり、個々の論文の引用数を表すものではないからだ。ごく一部の論文が例えば何百回も引用されて平均値を上げる一方、多くの論文の引用回数は少なく、分布を表す中央値は平均値よりかなり小さくなる場合が多い。ある雑誌が受けた引用の50%が、そこに掲載された論文の15%に集中していたという調査結果もある。掲載された雑誌のIFによって、個々の論文を評価し、さらには研究者の業績や能力を評価することに無理があることは明白だ。

分野によっても顕著な差がある。生物学系、とくに細胞生物学、生化学、分子生物学などでは高いが、解剖学や形態学、あるいは応用物理を除く理学ではおしなべて低い。研究者集団の大きさにも左右される。また、はやりのテーマに引用は集中しがちだ。

一方、IFの対象論文が過去2年間なのは、当初、IFへの関心が高かったのは分子生物学や生化学などで、これらの分野では引用対象は近年のものが多かったためという。だが、例えば、理論物理学などではさらに以前の論文が引用されることも多い。異なる分野の雑誌を比較することは適切でないと言われるゆえんだ。

警告は繰り返されてきた

こうしたことから、IFを考案したISI(現クラリベイト)のユージン・ガーフィールドをはじめ、同社は折に触れ、引用分析を研究者の業績評価に使わないよう注意を喚起してきた。私自身、1990年代末に米フィラデルフィアの同社取材で訪れた際、特にアジア諸国で評価に利用される例が多いことへの憂慮を聞いた。

有田正規・国立遺伝学研究所教授は雑誌「科学」の連載をまとめて今年10月に出版した「学術出版の来た道」(岩波書店)で、学術評価の第一人者であるガーフィールドの業績をたどり、「過去100年で3800万件の論文の半数は、自身も含めて1回も引用されず、10回以上引用される論文は2割に満たない」という2005年発表の「衝撃的な数字」を紹介している。有田さん

は、誰も引用しない研究に価値はあるのか、その社会的コストは、と問い、引用統計はトップよりむしろ末端を見るのに役立つと、としている。

「研究再現性の危機」が言われるまでに

本来、業績の評価には定性的な視点が不可欠のはずだ。しかし、客観的に見えることもあり、IFの数値が安易に使われてきた。90年代から注視してきた山崎茂明・愛知淑徳大名誉教授はそう語り、使う側と使われる側、その双方に責任があるとして、正しい理解に基づいて活用すべきだと訴えてきた。しかし、現実にはむしろ逆の方向に進んでいるようだという。

引用が多い傾向があるレビュー論文を増やしたり、IFの計算方法を逆手に取ったりして自誌のIFを上げようとする試みも後を絶たない。何より深刻なのは、IF至上主義の下、独創的な研究が生き残る余地が狭められる一方で、疑わしい論文が増え、特にこの10年来、生命科学を中心に「再現性の危機」が言われるまでになってきたことだ。

文部科学省の「政策評価から俯瞰するオープンサイエンス時代の研究評価の論点検討会」が7月にまとめた「新しい時代を見据えた研究開発評価の論点」¹⁾の中でも、次のように指摘されている。

「近年、定量的指標が過度に強調されることにより、研究不正や『研究再現性の危機』が生じるだけでなく、定量的評価では十分に評価されない萌芽の研究や時間をかけて深く掘り下げる研究などが忌避されるようになっていく」

文科省の科学技術・学術政策研究所の一線の研究者や有識者を対象とする定点調査でも、「IF偏重によって研究者の少ない分野は不利になりがちで研究分野の多様性が損なわれる」、「論文化しやすいテーマを安易に選ぶ若手が多い」といった声が上がっている。

me-too science が助長される

IFの偏重、つまり数値で測れるもので評価する風潮が多様性を損ない、本来あるべき科学研究の姿をゆがめていることが浮かび上がる。有力な学術誌の力の強まりとも裏腹の関係にあり、雑誌のオンライン化の流れとも相まって、容易な問題でないことも事実だ。

しかし、欧米を中心に、こうした現状を見直す動きがある。2012年には学術誌の編集者と出版者らが「研究評価に関するサンフランシスコ宣言 (DORA)」²⁾を公表し、科学的成果が賢明な方法で評価されることが大切だとして、助成機関や学術機関、研究者らに向けて勧告を行った。資金助成、職の任命や昇進に当たっ

て、IFのような定量的指標を排除し、雑誌ではなく研究自体の価値に基づいて評価することなどを挙げている。Science誌も、ブルース・アルバーツ編集長が、最も深刻な影響は流行を追う「me-too science」が助長されることとし、宣言を支持する論説³⁾を掲げた。日本からも日本生化学会などが名を連ねている。

また、2015年には研究評価における計量データの利用に関する「ライデン声明」⁴⁾がNature誌に発表された。オランダのライデン大学で開かれた科学技術指標に関する国際会議でのダイアナ・ヒックス博士(米ジョージア工科大)の基調講演をもとに科学計量学コミュニティがまとめたもので、「定量的評価は専門家による定性的評価の支援に用いるべき」、「機関、グループまたは研究者の研究目的に照らして業績を測定せよ」、「優れた地域的研究を保護せよ」、「指標を定期的に吟味し、改善せよ」など10の原則を掲げている。

測れるものを測るだけでない評価を

日本の動きはまだ鈍いようだが、決して手をこまぬいているわけにはいかない。言うまでもなく、今後の研究のあり方を左右するからだ。

目を転じれば、引用がトップ1%、10%などに入る日本からの注目論文はこのところ減る一方だ。IF重視の中、これはいったいどうしたことか。論文数自体が伸び悩んでいるとはいえ、IF偏重など評価の仕方に問題があったことを示唆してはいまいか。

研究評価は決して人任せにしていいものではない。受け身ではなく、研究者自身が研究を位置付け、評価した上で学界に問う。そこで独創性も主張できる。そうした評価までを含めたものが、研究ではないだろうか。そう心がけている研究者ももちろんいるだろうが、特にシニアの研究者には、IF頼みでなく、自らの研究の意義を堂々と問うことで、若い研究者に範を示してほしい。多様性に富んだ科学を育むため、測れるものを測るだけでない評価が、これからの日本の科学のためには不可欠だ。

日本化学会の見識とリーダーシップに期待したい。

- 1) https://www.mext.go.jp/content/20210730-mxt_kanseisk02-000017162_s6.pdf
- 2) <https://sfdora.org/read/>
- 3) <https://www.science.org/doi/epdf/10.1126/science.1240319>
- 4) http://www.leidenmanifesto.org/uploads/4/1/6/0/41603901/leiden_manifesto_japanese_161129.pdf

© 2021 The Chemical Society of Japan

ここに載せた論説は、日本化学会の論説委員会の委員の執筆によるもので、文責は基本的には執筆者にあります。日本化学会では、この内容が当会にとって重要な意見として掲載するものです。ご意見、ご感想を下記へお寄せ下さい。
論説委員会 E-mail: ronsetsu@chemistry.or.jp