



理化学研究所 生命医科学研究センター
センター長

東京大学 名誉教授

山本 一彦

【略歴】

昭和 27 年 神奈川県生まれ

昭和 52 年 東京大学医学部医学科 卒業

昭和 57

-60 年 ドイツ癌研究センター・

免疫遺伝学研究所 客員研究員

平成 7 年 九州大学生体防御医学研究所

臨床免疫学部門（内科学）教授

平成 9 年 東京大学大学院医学系研究科

アレルギー・リウマチ学 教授

平成 29 年 理化学研究所 生命医科学研究センター

副センター長

令和 2 年 現職

いろいろなスタイルの研究

生物学は、博物学の一部として発展してきましたが、いまや自然科学の中の大きな領域になっています。博物学は、観察、発見、分類などが主な方法ですが、その基本は今の生物学にも残っています。発見と帰納的考察という過程が中心と考えてよいかと思います。帰納的というのは、それまでのデータを矛盾なく説明できる包括的な結論を考えることです。個別な発見からそれを一般化するというとらえ方でも良いでしょう。さらに、18-19 世紀には、より新しい方法論として、仮説と演繹の科学が確立されました。これは既にあるデータなどをもとに論理的に推論し、問題点を明確にしてそれを解決するための仮説をたてる。この仮説が正しいか否かを試すために実験をおこなうというものです。一般化された考えを個別例で検証する、という考え方もできます。少数のサンプルを用いたメカニズム研究など、実験科学として必要な方法とも言えます。

歴史的経緯から、博物学の「発見と帰納」の科学は古い学問で、「仮説と演繹」が新しい学問だと考えられがちです。しかし、病原細菌の発見、ヒトゲノムの解読、機能的遺伝子のクロ

ーニングなど、今でもこの博物学的な生命科学の重要性は無くなっていません。ただし、多くの分野で「仮説と演繹」スタイルの研究がおこなわれてきた関係で、「仮説のない研究は研究に値しない」と極言されることもあります。

現在の研究のスタイルに当てはめてみると、仮説検証型研究と仮説生成型研究に分けることもできます。仮説検証型研究は、従来の情報を収集し考察し、仮説を検証するために適切な実験計画を立て実施します。統計学的手法などを用いたデータ解析を通じて、仮説が正しいかどうかに関しデータと仮説の一致度を判断し、仮説を棄却するか判定を保留するかを決めます。ただし、研究者が特定の仮説を持ちながら実験を行うため、結果に偏りが生じる可能性や、予想外の結果や薬剤に効果があるとは言えない結果が得られた場合、論文として執筆されなかったりすることがあります。雑誌に掲載する側の対応に関しても、特定のアプローチが優遇され、他の研究方法が排除される可能性があるとされています。これらのことから、時に権威主義的になりやすい方法論とも言われています。

科学的真実を明らかにする場合の基準として P 値が用いられることが多いのは周知の事実です。P 値が「統計的有意性」の境界を示す任意の閾値（例えば 0.05）を上回るか、下回るかで仮説を棄却するか否かが決まることとなります。しかしこの点で、「P 値」が看過できない“誤用”をもたらしているとの告発が 2016 年にアメリカ統計学会から出され、「統計的有意性」という用語を用いないことなどの提案がありました。P 値への過度の依存に対する懸念が表明されたものとされています。研究に用いるサンプルサイズが十分でない場合、有意差を含めた結果の信頼性が低下するなどの問題があり、逆に有意性を高めるために途中からサンプルサイズを増やすなどの不正が行われる事例も指摘されています (P-hacking)。そもそもこの統計推論に関する議論は、ネイマンとフィッシャーとの間で、立場や考え方の相違を含めて 20 年にも及ぶ論争が繰り返されたそうです。そして現在まで大方で受け入れられている「ネイマン・ピアソンの枠組みでの統計学」は、「意思決定のための理論」という前提でできた、かなり強圧的なものと言われています。一方、フィッシャーが目指した統計学的検定は、科学的帰納のための手段であり、データからいかに正確な推論を行なうかに主眼が置かれていました。

仮説検証型と違うアプローチの研究として、仮説生成型の研究があります。ある特定の仮説を証明するために実験を行うのではなく、バイアスのないデータを作り出し、そのデータのパターンや関係性を見つけて、新たな仮説を立てることを目的とする研究といっても良いでしょう。データ駆動型研究とも言われます。予想外の関連性や意外なパターンから、新たな仮説の芽生えや考え方が生まれる可能性や未知の現象への探求に向けた研究に繋がる可能性があります。この研究では、多数の変数に対して同時に検定が行われることから、P 値に替わり柔軟性の高い偽発見率 (false discovery rate: FDR) など使われます。次世代シークエ

ンサー、シングルセル解析など計測技術が進展する一方、統計学的解析だけでなく、人工頭脳を含めた解析手法の開発も進み、さらに最近の大規模データ学習による基盤モデルとトランスフォーマーの出現により、データ解析の可能性が広がっています。新たな「発見と帰納」の時代に入ったのではないのでしょうか。

そして、これらの研究手法は別々の方法論ではなく、仮説の設定には帰納法的思考が必要など、それぞれの局面でオーバーラップする場合があります。すなわち、これからの生命科学は多くの可能性を包含する科学が必要となると思われれます。新たな仮説が生成され、それが仮説検証型の研究の基盤となるなど、仮説生成型と仮説検証型の双方が科学の発展に寄与することが期待されます。そして、さらに柔軟な考え方の生命科学が必要となる時代がくるのではないのでしょうか。