

「統計的有意性やP値との付き合い方」

大阪医科大学 研究支援センター 医療統計室 室長・准教授

伊藤 ゆり



「P値が0.05より大きくなってしまったものは結果を示したくない」「 $P < 0.05$ で有意差が示せたので論文に書けそうです」など、P値の大きさに一喜一憂する人を数多く目の当たりにしてきました。2016年にアメリカ統計協会がこのような状況を憂慮し、「統計的有意性とP値に関する声明」を発表しました。この声明では以下のようなことが述べられています。

1. P値はデータと特定の統計モデル(仮説も含む)が矛盾する程度を示す指標の一つである。
2. P値は、調べている仮説が正しい確率や、データが偶然のみで得られた確率を図るものではない。
3. 科学的な結論や、ビジネス、政策における決定はP値がある値(有意水準0.05)を超えたかどうかのみに基づくべきではない。
4. 適正な推測のためには、すべてを報告する透明性が必要である。
5. P値や統計的有意性は、効果の大きさや結果の重要性を意味しない。
6. P値はそれだけでは統計モデルや仮説に関するエビデンスの、良い指標とはならない。

計量生物学会の声明の日本語訳より引用

つまり、 $P < 0.05$ は統計的有意に差があったのみを示しており、 $P < 0.05$ か否かで全てを判断・結論付けてはいけない、という警鐘です。例えば、統計的仮説検定を用いた新薬の臨床試験において、既存薬と新薬の生存率に統計的有意($P < 0.05$)に差があったとします(片側検定で対立仮説『新薬の生存率 > 既存薬の生存率』が採択された)。しかし、これにより直ちに新薬が既存薬に比べて優れているとは言えません。両群で副作用がどうであったが、どの程度の生存率の差があったかや脱落症例の理由、研究デザインは適切であったかなど、他にも多くの資料に基づき、

医学的に判断がなされる必要があります。「統計的有意差があった」だけで即OKではないのです。また「P値が0.05より大きかったので、二つの薬の効き目に差がなかった」とはならないことにも注意です。今回使用したデータでは帰無仮説が棄却されなかっただけであり、「差があるとはいえない」という程度の表現にとどめましょう。

統計の専門業界が声を上げているのは、P値だけで判断するのではなく、効果の大きさ自体や、信頼区間、統計モデルのよさの指標など様々な結果に基づいて、科学的な判断を下すようにしてほしいというメッセージです。いくつかの有名科学誌ではP値や「統計的有意性」という言葉の使用を禁じる流れもあります。2019年3月にはNature誌に800名を超える統計学者が「統計的有意性」という用語を使用しないことを提言しています。これまでのP値へのこだわりを捨て、「統計的有意差」だけで判断しないよう、様々な指標について確認し、科学的根拠の正しい報告の仕方を身につけていきたいですね。

参考文献

※1: アメリカ統計協会. 統計的有意性とP値に関するASA声明.

日本計量生物学会訳
<https://www.biometrics.gr.jp/news/all/ASA.pdf>

※2: Amrhein V, Greenland S, McShane B 他800名以上の賛同者. Scientists rise up against statistical significance. Nature. 20 March 2019.

<https://www.nature.com/articles/d41586-019-00857-9>

略歴

大阪大学大学院医学系研究科博士前期・後期課程卒業後、大阪府立成人病センター(現大阪国際がんセンター)リサーチ・レジデント、研究員、主任研究員を経て、大阪医科大学研究支援センター医療統計室准教授(現職)。現在、がん疫学、健康格差、医療統計の研究に主に従事。