

2025 年 7 月 17 日

要 望 書 1

「福島県県民健康調査」検討委員会 各位

永井宏幸

日本保健物理学会会員

理学博士

nagai.koko@gmail.com

私は放射線影響の疫学的研究を専門におこなっているものです。したがって統計学の学理と歴史についても理解を深める努力をしてきました。学理という語は、学問的論理、学問的道理の意味で使っています。

評価部会の「部会まとめ」には、統計学の学理に背く重大な問題があると考えております。

検討委員会には慎重な検討をしていただくよう要請いたします。

みなさんは、2019 年にNature誌に発表された「さらば 統計的有意 (Retire statistical

significance)」と題するエッセイを読んでいるのでしょうか。そこにはこう書かれています。

「統計学を学んでいない人には明確に見える差異を、統計学が科学者たちに否定するように仕向けるのはなぜなのだろうか？ 数世代にわたり、研究者たちは、統計的に有意でない結果は帰無仮説を証明するわけではないと警告されてきました」、「誇張した主張によるこうした誤解が学術文献を著しく歪めています。」

3人の統計学者に800人を越す研究者が連名して発表されたものです。ここには統計学の学理を知る研究者たちの危機感といらだちが現れています。

評価部会の「部会まとめ」は、福島県の小児甲状腺がんについて、「被ばく線量の増加に応じて発見率が上昇するといった一貫した関係は認められなかった」と結論しています。根拠に16枚のグラフが添えられています。第25回甲状腺検査評価部会資料2の図（A図とよぶ。）と参考資料2の図（B図と呼ぶ。）です。A図には4枚のグラフ、B図には12枚のグラフがあります。いずれも、被ばく線量を横軸にオッズ比を示したものです。グラフが右肩上がりになっていれば、がんの発見率が被ばく線量の増加に応じて上昇していることになります。16枚のグラフで14枚が右肩上がりになっています。避難地域に限定したB図の2枚のグラフだけは右肩上がりではありませんが、この地域の住民は避難経路によって被ばく状況が大きく変わるため、被ばく線量の推定が困難であることが知られています。したがって、全体としてみればグラフは

右肩上がりを示しています。ところが、評価部会のみなさんは、それは間違いだといいます。

統計学の教育を受けていないから右肩上がりだと思うのであって、統計学の教育を受けたわれわれは右肩上がりだと認めないといっているのです。

* * *

そこで、統計的有意とはなにかを、学理に沿って考えていきたいと思います。

統計的有意とは統計学の仮説検定で使われている専門用語です。仮説検定はこういうものです。まず仮説を立てます。これは帰無仮説とよぶことがあります。この仮説が正しいとして、その世界でつくられる標本の分布を考えます。データはこの標本のひとつです。そうしておいて、データが標本分布の中央から離れたところにあれば仮説とデータに**極端な食い違い**があると判定し、**仮説を棄却**します。極端でなければ仮説は棄却しません。極端かどうかはp値や信頼区間を使っての判定します。

仮説が棄却されたとき、これを**対立仮説が統計的有意になった**と表現することがあります。

対立仮説とは帰無仮説の余集合のことです。帰無仮説が「オッズ比は被ばく線量によって変化しない」であれば、対立仮説は「オッズ比は被ばく線量によって上昇する、または下降する」となります。仮説検定の学理からすれば、「対立仮説が統計的有意に認められた」という表現はともかくも、「対立仮説が統計的有意に認められなかった」という表現は不適切です。なぜなら、仮説検定はあくまで帰無仮説を検定するものであって、反対仮説を検定するものではないから

です。「対立仮説が統計的有意に認められなかった」という表現は、帰無仮説の正しさが認められたという意味合いをもってしまいます。じっさい、この結果から帰無仮説の正しさが認められたという結論を導いている科学論文が驚くほどあるのです。そのために、「さらば 統計的有意」は統計的有意ということばの使用をやめようといっているのです。統計的有意ということばをやめれば、統計学の誤用や誤解はほとんどなくなります。帰無仮説が棄却されなかったといえいいのです。それは、ほかの分析方法による分析やほかの集団の分析が必要であるということを書いていきます。物足りないかもしれませんが、これは仮説検定のもつ限界です。統計学には尤度比の方法やベイズ統計などのもっとすぐれた分析方法があります。

繰り返します。帰無仮説が棄却されなかったということは帰無仮説の正しさを示す証拠にはなりません。帰無仮説を棄却できないことを反対仮説が統計的有意でないと表現するのやめるべきです。そのような表現が、統計学の誤解や誤用を誘発し、社会に混乱を引き起こしています。

* * *

2016 年に「統計的有意とP値に関するASA声明」が発表されました。ASAは米国統計協会のことです。それからの引用です。

「研究者は、しばしばP値を帰無仮説が正しいという記述や、偶然の変動でデータが観察される確率に変えたがるが、P値はそのどちらでもない。比較的大きなP値であっても、帰無仮説

を支持するエビデンスとはならない。ほかのたくさんの仮説が、帰無仮説と同等か、それ以上に観察されたデータと矛盾しない可能性がある。」訳は日本計量生物学会国際担当理事の佐藤俊哉氏によるものです。ここでいうp値による仮説検定は信頼区間を利用した仮説検定と基本的に同じものです。

前段はすでに説明しているので、後段の「ほかのたくさんの仮説が、帰無仮説と同等か、それ以上に観察されたデータと矛盾しない可能性がある」について説明します。

仮説は帰無仮説のほかにもたくさんあります。いま、「線量に応じてオッズ比が右肩上がりに上昇する」という仮説を考えます。16枚のグラフのうち14枚は右肩上がりになっています。したがってこの仮説が棄却されないことはすぐわかります。「線量に応じてオッズ比が上昇する」という仮説は棄却できないのです。「部会まとめ」は片手落ちです。評価部会が公平な立場であるならば、「線量によって発生率は変化しない」という仮説は棄却できないが、「線量に応じて発生率が上昇する」という仮説も棄却できない、というべきなのです。

2013年7月28日の第21回評価部会で、鈴木部会長は「いずれの疫学的解析手法においても、被ばく線量と悪性ないし悪性疑い発見率との関連の解析において、被ばく線量の増加に応じて発見率が上昇するといった一貫した関係（線量・効果関係）は認められなかった」とするまとめ案を提示しましたが、祖父江委員はグラフの一枚を示し、「素直に上がってないとは言え

ない。有意差はないかもしれませんが、関連を示唆する結果であると捉えた方が自然だ」と発言してまとめ案に同意しませんでした。（議事録 25 頁を参照されたい。） 今回の部会まとめでは、「一貫した統計的有意が認められなかった」として、「一貫した」という限定が加わっています。一貫した統計的有意とは何かは説明していませんが 1 枚のグラフでは放射線の影響を認める条件にはならないといっているようです。14 枚すべてのグラフが統計的有意に上昇していなければ認めえないといっているのでしょうか。分析結果をみってから条件を釣り上げるのは公正な態度ではありません。

最後に、統計の教育を受けていない者にはどうして明白な差異が見えるのでしょうか。それはかれらが仮説検定の陥穽に囚われていないからです。

仮説検定では帰無仮説におおきなアドバンテージを与えます。帰無仮説を棄却するのは、仮説が 95%信頼区間の外にあるときに限っています。（p値でいえばp値が 0.025 より小さいときに限る。）これは非常に厳しい条件です。厳しくしているのは、本当は帰無仮説が正しいのに間違っ棄却してしまうことを避けたいからです。いま議論しているケースであれば、本当は放射線の影響がないのに影響があると判定することを避けたいからです。しかし、そのことによって、本当は放射線の影響があるのに影響がないと誤って判定する可能性が大きくなってしまいます。両者はトレードオフの関係にあるのです。

統計学を知らない者は特定の仮説にアドバンテージをつけたりしません。どの仮説も対等・公平に見ます。グラフを見て、右肩上がりになっているかそれとも水平に近いかを素直に判断します。統計学を知らない者の判断が仮説検定にもとづく専門家の判断と大きく食い違うのはこのためです。おもしろいことに、統計学を知らない者の判断は、尤度比やベイズ統計を用いた判断との親和性が非常に高いのです。

統計学を知らない市民の方が、学理を知らない統計の専門家よりも本当のことがわかっているのです。奇妙な光景です。ある人はこの状況を「科学界における汚れた秘密(Dirty secret)」と形容しています。ある人は「王様は裸だ」と笑っています。福島の小児甲状腺がんの疫学研究が統計学の正道に沿ってすすむことを願ってこの要望書をお渡しいたします。

引用文献

1. "Retire statistical significance," Valentin Amrhein, Sander Greenland, Blake McShane and more than 800 signatories 2019, DOI:10.1038/d41586-019-00857-9, <https://www.nature.com/articles/d41586-019-00857-9>
2. "The ASA's statement on p-values: Context, process, and purpose", Wasserstein & Nicole A. Lazar, The American Statistician 2016; 70: 129-133.
<http://dx.doi.org/10.1080/00031305.2016.1154108>.

邦訳：「統計的有意性と P 値に関する ASA 声明」, 佐藤俊哉, 2017 年,

<https://biometrics.gr.jp/news/all/ASA.pdf>.

3. 第 21 回甲状腺検査評価部会議事録（2023 年 7 月 28 日開催） p.25-.

