

名古屋市HPVワクチン接種後調査データを用いた2つの解析論文の比較

八重ゆかり
聖路加国際大学

Comparison between two papers analyzing data from the “Nagoya City Cervical Cancer Immunization Program Survey Results”

Yukari Yaju
St. Luke’s International University

Abstract

A self-completed mail survey of adverse events experienced in approximately 71,000 young women, the “Nagoya City Cervical Cancer Immunization Program Survey,” was conducted in September 2015. The summary and the complete questionnaire data of the Nagoya Survey are available at the website of Nagoya City. Dr. Suzuki et al. stated that none of the reported symptoms were significantly associated with increased odds of them occurring after the administration of the HPV vaccine based on the results of the Nagoya Survey in their article published in the journal “Papillomavirus Research” in 2018. Conversely, the author and Dr. Tsubaki re-analyzed the same data and reported that a possible association between HPV vaccination and some symptoms exists in the paper presented in the journal “Japan Journal of Nursing Science” in 2019. Dr. Suzuki has raised some concerns in a Letter to the Editor and commented on some issues. After thorough consideration of the Letter to the Editor from Dr. Suzuki, the author and Dr. Tsubaki answered each of his points in the Authors’ response. However, Dr. Suzuki made the same comments in the Letter to the Editor as well as in the Asahi Shinbun speech website “Web Ronza” and the journal “Pediatrics of Japan.” In this article, an outline of the process of discussion is shown in Section 1, and details are described in the later section. The author sincerely believes that this article will stimulate academic discussion and contribute to a better understanding of HPV vaccine safety.

Key words

Nagoya City, human papillomavirus (HPV), cervical cancer, vaccine, statistical analysis

Rinsho Hyoka (Clinical Evaluation). 2022 ; 49(3) : 483-98.

1. はじめに

2015年9月、約7万1千人の若年女性を対象とした自記式質問紙によるヒトパピローマウイルス (human papillomavirus, 以下HPV) ワクチン接種関連の有害事象に関する調査「名古屋市子宮頸がん予防接種調査」(以下、名古屋調査) が実施された。調査の理由として名古屋市は、副反応を訴える方々から実態調査を求める声が寄せられたことから実態を把握する必要があると考え実施したとしている¹⁾。名古屋市は「名古屋市子宮頸がん予防接種調査解析結果(速報)」(以下、速報) を名古屋市ホームページ公開したが、その後この速報は公開が取りやめになり、現在名古屋市では、調査の全体集計結果と調査データをホームページに公開している¹⁾。2018年、鈴木と細野は、名古屋調査データを分析したところHPVワクチン接種との関連が認められた症状はなかったとする結果を、英文誌 *Papillomavirus Research* で報告した²⁾ (以下、鈴木・細野論文)。一方2019年、筆者と椿は名古屋調査データを用い、異なる解析手法を用いて検討した結果、調査対象となった24症状のうち、HPVワクチン接種後症状として報告されているいくつかの特徴的症状について接種との関連の可能性が示唆されたことを、日本看護科学学会発行の英文誌 *Japan Journal of Nursing Science (JJNS)* にて報告した³⁾ (以下、八重・椿論文)。

その後、鈴木氏から *JJNS* 編集部あてに八重・椿論文に対する批判レターが投稿され公表された⁴⁾ (以下、鈴木レター)。これに対する八重・椿からの著者回答も *JJNS* において公表されている⁵⁾ (以下、著者レター)。また鈴木レターに対する *JJNS* 編集長からの回答もレターとして公開されている⁶⁾。

筆者は、前記の著者レター公表をもって鈴木氏による批判に対する学術的回答は完了していると考えているが、鈴木氏は *JJNS* レターでのやり取りの後、様々なメディア等を通して八重・椿論文への批判を公表している^{7, 8, 9)}。

本稿ではまず、鈴木・細野論文と八重・椿論文のそれぞれの結果の概要を紹介したうえで、*JJNS* レターにおける鈴木氏からの八重・椿論文に対する指摘事項とそれらに対する著者回答の概要を、6つの論点として紹介する。つぎに、*JJNS* レターでのやりとりでの議論内容を基本としつつ、その後、朝日新聞言論ウェブサイト「論座」^{7, 8)} (以下、朝日論座)、及び雑誌「小児科」⁹⁾ (以下、雑誌小児科) で行われた鈴木氏による追加の主張も踏まえ、現時点における筆者の見解を詳述する。なお、朝日論座と雑誌小児科に示された鈴木氏の主張には「*JJNS* レターで議論された内容が繰り返されている部分」と「新たな主張の部分」とが混在しているため、どちらに対する筆者の見解を述べているのが区別できるような表現を心掛けた。さらに鈴木・細野論文と八重・椿論文の結果の違いに端を発した論争において特に重要と考えられた、検定の多重性の調整について追加的解説を行った。

2. 鈴木・細野論文と八重・椿論文について

2.1 鈴木・細野論文、八重・椿論文、それぞれの結果・結論の概要

鈴木・細野論文と八重・椿論文の主要な結果と結論を紹介する。各論文の内容は直接、原論文を参照されたい。

鈴木・細野論文²⁾ では主な結果として以下の内容が提示され、「HPVワクチン接種と24症状発現との間で統計学的有意な関連は認められず、このことから、ワクチンと報告されている有害事象との関連はないことが示唆された(翻訳、筆者)」との結論が導かれている。

- 24症状の発現を primary outcome として定義。
- 接種と年齢を共変量としたロジスティック回帰分析(複数の要因/曝露[説明変数]から二値の結果[反応変数; 疾患発生の有無など]が起こる確率を説明・予測するために用いられる回帰分析であり、曝露の影響の指標とし

てオッズ比が示される、多変量解析の1種である。)を行い、接種の影響(症状発現)の年齢調整オッズ比を提示。24症状中5症状ではオッズ比が1以上であるが95%信頼区間が1をまたいでおり統計学的有意ではなかった。残り19症状ではオッズ比が1未満であり、うち14症状では統計学的有意、5症状では有意ではなかった。

- 調査時年齢、初回接種年齢、接種年の組み合わせによるコホート1～5を設定し、サブグループ解析を実施。サブグループの抽出条件は以下。

コホート1：接種群・非接種群ともに調査時年齢20-21歳、接種群の初回接種年齢16-17歳、接種年2011年

コホート2：接種群・非接種群ともに調査時年齢18-19歳、接種群の初回接種年齢13-14歳、接種年2010年

コホート3：接種群・非接種群ともに調査時年齢17-18歳、接種群の初回接種年齢13-14歳、接種年2011年

コホート4：接種群・非接種群ともに調査時年齢16-17歳、接種群の初回接種年齢13-14歳、接種年2012年

コホート5：接種群・非接種群ともに調査時年齢15-16歳、接種群の初回接種年齢13-14歳、接種年2013年

- サブグループ解析の結果(年齢調整オッズ比)において、コホート4(接種群・非接種群ともに調査時年齢16-17歳、接種群の初回接種年齢13-14歳、接種年2012年)では「突然力が抜ける」で、またコホート5(接種群・非接種群ともに調査時年齢15-16歳、接種群の初回接種年齢13-14歳、接種年2013年)では「物覚えが悪くなった」でオッズ比が1を上回り、統計学的有意であった。

一方、八重・椿論文³⁾では複数の探索的解析を実施し、主として以下の結果を提示したうえで、“HPVワクチンと認知機能障害、運動機能障害な

どの特徴的な症状との関連の可能性が示された”との結論を導いている。

- 接種と年齢を共変量としたロジスティック回帰分析を行い、症状発現の年齢調整オッズ比を提示。24症状中2症状ではオッズ比が1以上であるが95%信頼区間が1をまたいでおり統計学的有意ではなく、オッズ比が1未満となった残り22症状のうち4症状では統計学的有意ではなく、18症状では統計学的有意であった。
- 接種とStudy period (Study periodについては後述)を共変量としたロジスティック回帰分析を行い、症状発現のStudy period調整オッズ比を提示。24症状中7症状でオッズ比が1以上となり、うち2症状(「物覚えが悪くなった」,「身体が自分の意思に反して動く」)で統計学的有意であった。
- 接種と年齢、接種とStudy period、それぞれの交互作用の検定結果を提示。ほとんどの症状において統計学的有意な結果となった。
- 接種、Study period、及び接種とStudy periodの交互作用を組み込んだモデルでのロジスティック回帰分析を行い、症状発現のオッズ比を提示。24症状中4症状(「物覚えが悪くなった」,「簡単な計算ができなくなった」,「簡単な漢字が思い出せなくなった」,「突然力が抜ける」)でオッズ比は1を上回り統計学的有意な結果であり、交互作用の検定も統計学的有意であった。
- 年齢層別解析結果を提示。15-16歳層での解析の場合、3症状(「物覚えが悪くなった」,「簡単な計算ができなくなった」,「身体が自分の意思に反して動く」)で症状発現の粗オッズ比が1を上回り統計学的有意な結果であった。
- 健康者接種バイアスの影響に対処する方法として、全年齢層の接種群と15-16歳年齢層の非接種群との比較を実施。症状発現の粗オッズ比は「物覚えが悪くなった」,「簡単な計算ができなくなった」,「簡単な漢字が思い出せ

なくなった」,「身体が自分の意思に反して動く」,「普通に歩けなくなった」,「突然力が抜ける」の6症状で1を上回り統計学的有意な結果であった。

2.2 JJNSレターでの鈴木氏による八重・椿論文に対する指摘事項と八重・椿からの回答概要：6つの論点について

JJNSレターにおいて鈴木氏から指摘された事項、及び八重・椿からの著者回答における主な論点は、Study period, ロジスティック回帰モデルでの年齢調整, 交互作用, 年齢層別解析における対照群の設定, 検定の多重性の調整, 利益相反の開示の6点であった。鈴木氏からの指摘事項と著者回答の概要をTable 1にまとめた。

3. JJNSレターでの6つの論点に関する、その後の経過を含めた筆者の見解

以下では、Table 1に概要を示したJJNSレターでの議論の項目ごとに、議論内容の詳細を紹介するとともに、その後、朝日論座及び雑誌小児科で行われた鈴木氏による追加の主張も紹介しつつ、それらを踏まえたうえで、現時点における筆者の見解を記述する。

3.1 Study period

鈴木レターでは、“八重・椿論文で用いたStudy periodのデータは、接種群では初回接種時点を起点としており、この場合非接種群の起点は初回接種するはずであった時点でなければならないが、八重・椿論文では調査期間の開始時点である12歳を起点としているため、接種者と非接種者のStudy periodはCounterfactualの前提を満たしておらずStudy periodを用いた解析は誤った結果を生み出す”と指摘されている。まず、八重・椿論文がStudy periodを採用した理由は以下のとおりである。一つは年齢よりもStudy periodのほうが「症状経験を思い出す期間」としてより妥当であると考えられたことである。もう1点は、ロジス

ティック回帰モデルで年齢を共変量とした場合には交互作用の存在が認められたため、接種と年齢を共変量とする主効果だけのモデルは不適切であると考えられ、さらに年齢に代わる他の変数の可能性を探る必要性があると考えられたことである。

さて、3.2ロジスティック回帰モデルでの年齢調整、及び3.3交互作用の項でも後述するが、筆者が24症状について年齢層ごとのイベント発生割合を確認したところ、年齢増加とともに症状経験割合が一樣に増加または減少するという傾向が、接種群・非接種群に共通して認められた症状は「足が冷たい」(一樣に増加傾向)のみであり、他の23症状においては、一樣に増加または減少という傾向が両群ともに認められない、あるいは少なくともどちらかの群で認められないという状況が観察された。また若年齢群(たとえば15-16歳年齢層など)と高年齢群(たとえば18歳以上年齢層など)とでは症状経験割合に明らかな違いが存在する状況も認められ、若年齢群と高年齢群では何らかの集団特性が異なっている可能性も示唆された(八重・椿論文では、これら年齢とイベント発生割合との関係性に関する検討結果を示しており、代表例として「物覚えが悪くなった」での結果を図1~3、として提示している³⁾。これらの知見から、接種と年齢との間には交互作用が存在すると考えられ、接種と年齢との交互作用の検定(ロジスティック回帰モデルでの交互作用の検定)を行ったところほとんどの症状(24症状中18症状)において統計的有意の結果となった³⁾。このような経緯を踏まえ、筆者はロジスティック回帰分析を用いる場合には、年齢とは異なる調整変数の可能性も探る必要性があると考え、Study periodを設定し調整変数とした解析を試みたものである。名古屋調査では調査対象者(15歳から21歳)が調査対象期間(12歳時点から2015年の調査時点の年齢時まで)においてHPVワクチン接種後症状の可能性が指摘されている当該症状を経験したかどうかについてのデータ収集をしている。そしてこのStudy periodは、「症状経験を思

Table 1 Outline of the discussion in the *JJNS* Letters^{4, 5)}

項目	鈴木氏の指摘事項	八重・椿(著者ら)からの回答
Study period	Study periodは接種群と非接種群で系統的に異なっており、接種群と非接種群の比較がCounterfactualの前提を満たしていないため、このようなStudy periodの使用は不適切である。	Study periodは「症状経験を思い出す期間」として設定したものである。年齢を「症状経験を思い出す期間」として用いることも可能であったが、Study periodのほうがより適切な期間と考えたため解析における調整変数として用いたものであり、探索的解析方法の1つとして許容されると考える。なお接種群と非接種群でのStudy periodの長さの系統的違い(接種群のほうが短い傾向)の存在については鈴木氏のご指摘のとおりである。ただし、接種群のほうが短い傾向にあったことから、この系統的違いが症状経験オッズ比の推定値を上昇させる方向に働いたとは必ずしも言えないと考える。
ロジスティック回帰モデルでの年齢調整	現象としての交絡がある(単変量解析と年齢調整解析の結果が異なる)場合は、交絡の3条件は満たしているはずであり、名古屋調査データでは年齢による交絡が存在することは明らかで、年齢による交絡が観察されているにも関わらず、年齢で補正しない結果を提示することは誤りである。	上記でも述べたとおり、調査対象者の年齢については、生物学的年齢ではなく「症状経験を思い出す期間」の長さを反映する値としてとらえる場合には調整因子として用いることは妥当であると考えたため、著者らは粗オッズ比とともに、ロジスティック回帰分析による年齢調整オッズ比も提示している。したがって鈴木氏の指摘は当たらない。
交互作用	交互作用項を解析に入れるのは交互作用に興味の焦点があるときであり、交互作用項を加えた解析での主効果については、慎重な考察が必要である。八重・椿論文では、交互作用項を回帰分析に用いた結果を示すことで、ミスリーディングな主効果を提示している。	年齢層別及びStudy period層別の症状発現割合をみることにより、接種と年齢、接種とStudy period、どちらの間にも交互作用(交絡変数として複数の変数を想定した場合に、アウトカム変数に対してある変数と他の変数が相乗的に影響すること ²¹⁾)の存在が想定されたため、著者らは交互作用の検定を行った。ほとんどの症状において交互作用の存在が示唆された。そしてほとんどの症状において量的交互作用(交絡変数をサブグループに分けた場合に、どのサブグループでも一貫して因果的[または逆に予防的]な効果があるが効果の大きさは一定していないこと ¹⁴⁾)が存在する。いくつかの症状では質的交互作用(交絡変数をサブグループに分けた場合に、あるサブグループでは因果的な効果があり、別なサブグループでは予防的な効果があるように効果の方向が逆転してしまうこと ¹⁴⁾)が存在することを確認した。そこでロジスティック回帰モデルでは、接種の有無、Study periodの他に、接種とStudy periodとの交互作用項を変数に加えたモデルも検討すべきと考えて結果を提示したものである。著者らは、交互作用が存在する場合に、交互作用を考慮しない解析による主効果を提示することのほうがミスリーディングな結果の提示であると考えた。
年齢層別解析における対照群の設定	15-16歳年齢層の非接種群を対照群とした全年齢層接種群のオッズ比を示しているが不適切な比較であり、15-16歳年齢層の非接種群を対照群とする場合には15-16歳の接種群のみを使用するべきである。	15-16歳の非接種群を対照群とした全年齢層接種群のオッズ比は、健康者接種バイアスの影響をできるだけ少なくすることを目的とした解析方法として行ったものである。名古屋調査の15-16歳の非接種群は、健康状態不良のために接種できなかったというよりも健康でかつ接種しなかった人という想定が可能と考えられた(根拠の詳細は本文参照)。したがって健康で接種が可能であった全年齢層接種群に対して、15-16歳非接種群を対照として設定することにより、健康者接種バイアスの影響が比較的少ない状態での比較が可能となるのであり、このような解析は探索的解析手法の一つとして妥当であると考えた。
検定の多重性の調整	24症状について検定を行っているため検定の多重性の調整が必要である。	著者らは、安全性の検討においては検定の多重性(ある命題の真偽を判定するために複数の[多くの]統計的推測[とくに検定]を行う(重ねる)こと、検定における第1種の過誤率または/かつ第2種の過誤率が制御されないこと ²²⁾)の調整は不要と考える。著者らは統計解析の項に多重性の調整を行わない旨を明言した上でこれを実施せず、多重性の調整を行わない結果を提示しており適切な対処方法と考える。
利益相反の開示	筆頭著者の八重は、薬害オンブズパースン会議のメンバーであったにも関わらず、論文にはCOI事項としてこのことが明記されておらず不適切である。	論文上でのCOI報告については、その手続きに不備はないと考える。八重は投稿時、当時薬害オンブズパースン会議のメンバーであったことを申告しており、この事実が論文中に記載されなかった理由は、雑誌 <i>JJNS</i> の規定によるものである。このことは編集長からのレターでも明確に説明されている。

い出す期間」あるいは「症状経験が可能であった期間」の値として、接種群は「初回接種時点から2015年時点までの年数」、非接種群は「12歳時点から2015年時点までの年数」として定義したものである。

そこでStudy periodについてであるが、八重・椿論文におけるStudy periodの範囲は、ワクチン接種群が0-6年、非接種群は3-9年であり、このStudy periodの長さにおける両群間での系統的違い（データ範囲が異なり、かつ平均的に接種群のほうが短い傾向にある）が存在することは、Table 1でも述べたとおり鈴木氏の指摘のとおりである。ただしStudy periodの長さにおける両群間での違いは、症状発現割合を見る際に、接種群のほうが非接種群よりも相対的に低く見積もられる（「症状経験を思い出す期間」が短いことにより、経験回数が低く見積もられる）可能性があることを考えると、結果として症状経験オッズ比を低下させる要因となった可能性があるのではないだろうか。また、この系統的違いが症状経験オッズ比の推定値を上昇させる方向に働いたとは必ずしも言えないのではないかと考える。八重・椿論文では、このような条件下においてさえ、非接種群に対する接種群での症状発現オッズ比が1を上回る症状（「物覚えが悪くなった」、「簡単な計算ができなくなった」、「簡単な漢字が思い出せなくなった」、「突然力が抜ける」、など）が認められたことには、安全性に関するシグナル検出としての意義があると考えた。ただし、このStudy periodを共変量とするロジスティック回帰モデルについては、接種とStudy periodとの交互作用が有意であった症状が、年齢の場合よりも少ない数ではあるものの存在しており（24症状中7症状）、八重・椿論文では解析手法としては課題が残るものであることを明記したうえで結果を提示している³⁾。なお、交互作用が存在することに基づく問題、また交互作用をモデルに組み込むうえでのStudy periodに関連した問題については3.3交互作用の項で詳述する。

鈴木氏が主張する、Counterfactualな対照群と

の比較が理想的であるという点に関しては、筆者も同意するところである¹⁰⁾。すなわち、Study periodを用いた場合に、接種群に対して非接種群がCounterfactualな対照群にはならないとの鈴木氏の指摘はそのとおりである。そこでこのようなCounterfactualな対照群との比較の可能性を検討するとすれば、たとえば接種時年齢マッチングの方法が一つの代替案として考えられるであろう。ただし、この手法には健康者接種バイアスを考慮できないという限界があることには注意が必要である。他に、自己対照ケース・シリーズデザインも代替案の一つと考えられる。しかしながら、HPVワクチン接種後症状の発現が接種後数カ月から数年という長期にわたる可能性が指摘されていること、名古屋調査では複数回経験した症状であっても初回時点の情報しか得られていないことから、この自己対照ケース・シリーズデザインは名古屋調査データの解析には適していないと考えられる。

3.2 ロジスティック回帰モデルでの年齢調整

鈴木レターでは“現象としての交絡がある（単変量解析と年齢調整析の結果が異なる）場合は、交絡の3条件は満たしているはずであり、名古屋調査データでは年齢による交絡が存在することは明らかであるから、年齢による交絡が観察されているにも関わらず年齢で調整しない結果を提示することは誤りである”，との指摘がなされている。また、朝日論座や雑誌小児科においても、“八重・椿論文では年齢調整をしないことが正当化されている”，“年齢調整の概念が理解できていない”，との指摘もなされている。

八重・椿論文ではまず、年齢を交絡因子として扱う前提としては、生物学的年齢が症状発症に関連し得るという仮定が必要であるところ、15-21歳という調査対象年齢層において検討対象となった24症状が、年齢増加とともに発症しやすくなるという仮定は医学的に考えにくく、年齢を交絡因子として扱うことは必ずしも妥当とは言えないと考えた。もちろん疾患によっては、若年層と成

人とで発症割合が異なる、あるいは十代の中でも年代により発症割合が異なる場合も存在するであろう。しかし名古屋調査で対象としているのは何らかの疾患として診断されるような既存疾患ではなく、あくまでもHPVワクチン接種後症状の可能性があると考えられた個別の24症状であり、八重・椿論文ではこれら24症状が15～21歳という年齢層の中での年齢上昇とともに発症しやすくなるという前提を置くことの妥当性は低いと考えたものである。

ただし3.1 Study periodの項でも述べたとおり、八重・椿論文では、名古屋調査の年齢データを、生物学的年齢ではなく「症状経験を思い出す期間」として捉えることは可能であると考え、ロジスティック回帰分析による年齢調整結果も提示している。すなわち、八重・椿論文では年齢データでの調整を全く否定しているわけではないことをまず確認しておきたい。その上で接種と年齢との交互作用の存在の観点から、年齢を共変量としたロジスティック回帰モデルを用いることには問題があると指摘している（交互作用の問題については3.3 交互作用の項で詳述する）。また八重・椿論文では、年齢層別での粗オッズ比を提示することにより、接種と症状発現との関連において、年齢がどのように影響しているかの分析を行っている。すなわち年齢データでの調整という考え方を否定しているわけではなく、年齢層の違い（年齢が異なることによる集団特性の違い）を踏まえたうえで、より適切な解析方法は何かを検討し結果を提示している。

なお鈴木ライターでは、“年齢は標準的交絡因子であり、疫学的解析において年齢を交絡因子として用いるのが当然である”との主張もなされている。疫学研究の解析において年齢を交絡因子として扱うことについては、一般論としては筆者も同意するところである。たとえば木原らは、“年齢は多くの健康アウトカムの重要なリスクファクターであり、かつ他の多くの要因への曝露とも関係があります。したがって、年齢を主たる目的とする研究でなくても、年齢が交絡因子Confounding

factorとなる可能性があるため、年齢と曝露、そして年齢とアウトカムの関係は常に注意深く検討しなければなりません。”と指摘している¹¹⁾。しかし、何の値を交絡因子として扱うべきか、モデルの共変量として組み込む値としてはどの値が妥当かについては、研究対象とするデータの詳細かつ丁寧な検討に基づいて決定されなければならない。たとえば山口は、ロジスティック回帰分析に関する解説記事の中で、“変数選択には、(中略)いわゆる統計学的変数選択と医学的知識を利用した選択がある。”と述べたうえで医学的変数選択の方法を説明、推奨するとして、“①曝露変数は必ずモデルに入れる、②交絡要因であることがわかっている変数は必ずモデルに入れる、③強い予後因子（結果を予測する因子）は必ずモデルに入れる、④曝露と結果とのあいだの中間変数はモデルに入れてはいけない、を原則とする。さらに、交絡要因の可能性のある変数については、前述したとおり、できるだけ多くの変数をモデルに入れると曝露効果の推定のバイアスが小さくなる一方で、不必要な変数をたくさんモデルに入れると曝露効果の推定誤差が大きくなり信頼区間が広がるため、交絡要因の可能性のある変数だけを変数選択の対象とする”という方法を提唱している¹²⁾。単に、疫学における標準的な交絡因子であるから交絡因子として扱う、あるいは回帰モデルの共変量に組み込むことが当然であるとは言えないのである。

3.3 交互作用

ここではまず、交互作用が存在する場合に、交互作用を考慮しない結果により主効果を主張することの問題点について確認しておきたい。交互作用が存在するデータにおいてその交互作用を無視した解析により主効果を主張すると、交互作用の存在が検証された場合に主効果の推定をしてはならないという統計学的原則に反することとなる。つまり解析対象とするデータにおいて、まずは交互作用の有無について予備的検定を行い、主効果の検定のためのモデルに交互作用を入れるかどうか

かを決定すること、その後主効果の検定を行い、有意差が認められる場合に主効果があると判断する、このような手続きを経るのが統計解析方法の原則である。丹後は著書「統計学のセンス」の中で、“交互作用を無視して（交互作用に気づかずに）交絡因子の調整の計算を行うとその結果は明らかに誤りである。”と指摘している¹³⁾。

また佐藤は「疫学研究における交絡と効果の修飾」と題する1994年の論文¹⁴⁾において次のように指摘している：“交絡の調整は効果の修飾 effect modification が存在するかどうかで方法が異なる。効果の修飾は、効果の指標と交絡変数との（統計的）交互作用に相当するが、交互作用を量的と質的に分けた場合、量的な交互作用のことを指す（Miettinen (1974)¹⁵⁾）。量的な交互作用とはどのサブグループでも一貫して因果的（または逆に予防的）な効果があるが効果の大きさは一定していないことであり、質的な交互作用とはあるサブグループでは因果的な効果があり、別なサブグループでは予防的な効果があるように効果の方向が逆転してしまうことである。効果の修飾がない場合には回帰モデルなどを用いて交絡を調整した共通な効果の指標を求めることができる。（下線は筆者）”。この佐藤の指摘を名古屋調査データに当てはめると、効果の指標（HPVワクチン接種の影響としての症状発現のオッズ比）と交絡変数（年齢）との交互作用がない場合に回帰モデルでの交絡変数（年齢）の調整を行うことができる、ということになる。つまり、交互作用の存在が確認された場合には回帰モデルでの主効果の推定をしてはならないのである。

特に、質的な交互作用がある（説明変数のある層までとそれ以降の層で反応変数への影響が逆転する）場合には、交互作用を考慮しないロジスティック回帰モデルでの調整結果を提示することは明らかに統計的方法の誤用である。この質的な交互作用に関して、また交互作用に関する検討と解析方法の探索の重要性については、前述の佐藤論文では以下のとおり指摘している：“質的な交互作用が存在する場合にはもはや結果を単一の効果

の指標に要約することはできない。したがって、（中略）重大な交互作用を見落とさないために「各層ごとの効果の指標のチェック」をおこなってはならない。また、疫学研究では曝露-疾病発生間の真実のモデルが分かっている例はほとんどないので（Greenland (1989b)¹⁶⁾）、現在のデータ解析ではCox回帰、ロジスティック回帰がルーチンとなっているが、これら以外のモデル当てはめの検討、層別解析の併用による結果の確認、なども必ず実施しなければならない”。すなわち、量的または質的な交互作用の存在を見落とさないためのデータの探索を丁寧に行うこと、交互作用の存在が確認された場合には他のモデルや別の解析方法の検討を行うことが重要であり、少なくとも、交互作用が確認された交絡変数を用いたモデルでの調整結果で主効果を主張してはならないということを佐藤は指摘している。以上の交互作用の問題を名古屋調査データに当てはめて考えるならば、年齢調整のためにロジスティック回帰モデルを用いる場合には、接種による影響の指標（症状発現）と年齢の間には量的、及び質的な交互作用が存在するという事実があることを忘れてはならないのである。

さて鈴木レターでは、“交互作用項を解析に入れるのは交互作用に興味の焦点があるときであり、交互作用項を加えた解析での主効果については、慎重な考察が必要である”と指摘している。その後鈴木氏は、八重・椿論文が交互作用について検討し交互作用を考慮した結果を提示していることに対し、朝日論座や雑誌小児科において“交互作用存在下での主効果を示すことにより主効果に誤解を招いている”、“ワクチンの効果を恣意的に論じている”と主張し、名古屋調査データでの交互作用の検討の必要性を否定しているともとれる論を展開している。交互作用を組み込んだ解析での主効果について、慎重な解釈が求められることは鈴木氏の主張のとおりであるが、交互作用の検討をしなくてよいわけではなく、交互作用の存在が認められる場合に交互作用を無視した解析を行うと結果は間違っただけのものとなることに注意しな

ければならない。

八重・椿論文ではまず、接種群、非接種群それぞれにおいて、年齢層別、またStudy period層別のイベント発生割合のヒストグラムを24症状全てに作成し、年齢層またはStudy period層ごとのイベント発生割合状況を確認した。さらに、年齢に関しては質的、量的交互作用についても検定し、著者レターでは「Table 0 Results of test for interaction between human papilloma virus vaccination and age (HPVワクチン接種と年齢との交互作用に関する検定結果)」も提示している。以上の知見から言えることは、①年齢またはStudy periodを共変量としたロジスティック回帰モデルでの調整オッズ比の値は不適切であると考えられること、また②接種と年齢または接種とStudy period、それぞれの交互作用項を組み込んだロジスティック回帰モデルでの調整オッズ比の値についてもその妥当性には限界があることである。

なおStudy periodについては、3.1で言及した、長さにおける接種群と非接種群での系統的違い（データ範囲が異なり、かつ平均的に接種群のほうが短い）と関連して、ロジスティック回帰モデルで変数として用いた場合の問題点についても言及しておきたい。八重・椿論文ではTable 4に、接種とStudy periodの交互作用項を組み込んだロジスティック回帰モデルでの調整オッズ比を提示している。この場合、Study periodは接種群のデータ範囲が0-6年、非接種群では3-9年であることからモデル解析では基準点が0（非接種群では存在しない値）となり、推定された値が特別な層の主効果を際立たせた結果になった可能性がある（たとえば「物覚えが悪くなった」のオッズ比が3.59など、いくつかの症状でオッズ比が特に大きい値として推定されている）があることに注意が必要である。そして、このような値を接種と症状との因果関係を最終的に推論する値として扱うことは適切でなく、慎重な解釈が求められることは言うまでもない。これらの点については八重・椿論文で言及できていなかったこと、また、使用するStudy periodのデータ範囲を変えた感度解析

を実施せずに全データだけでの解析結果だけを提示したことは、八重・椿論文の結果提示や説明において不備であった点と考える。

そこで今回、Study periodのデータ範囲を変えた感度解析を実施したのでTable 2に結果を提示する。以下の3種類のStudy periodを設定した解析を接種、Study period、及び接種とStudy periodの交互作用項を組み込んだロジスティック回帰モデルで実施し、症状発現のオッズ比を算出した結果である。なおTable 2には八重・椿論文に提示した接種群0-6年、非接種群3-9年のデータセット（Study period 0～9年）の結果を比較のために転載して示した。

- 接種群0-6年、非接種群3-6年のデータセット（Study period 0～6年）
- 接種群3-6年、非接種群3-6年のデータセット（Study period 3～6年）
- 接種群3-6年、非接種群3-9年のデータセット（Study period 3～9年）

Table 2に示すとおり、感度解析の結果では、八重・椿論文での結果（たとえば「物覚えが悪くなった」でオッズ比は1を上回り統計学的有意であり、交互作用の検定も有意であった）と類似した結果が再現された。八重・椿論文では、接種とStudy periodの交互作用項を組み込んだロジスティック回帰モデルでの調整オッズ比は、あくまでも探索的解析の経過を丁寧に表示す必要性を考え、またモデル解析上の問題を含む値であったとしても、危険性のシグナル検出結果の1つとして提示することが否定されるものではないとの判断のもとに示したものである。しかし、提示した値が因果推論を結論づける値であるかのように利用される可能性を考慮するならば、今回のような感度解析結果を伴うことなく、議論のあり得る値だけを結果として提示したことがミスリーディングであったとする指摘は妥当であり、八重・椿論文での結果の提示のしかたでの不備な点であったと考える。なお八重・椿論文では、年齢を調整する方法としてStandardization法についても検討し考察に、量的交互作用を考慮した解析結果の一例

Table 2 Sensitivity analysis; Study period-adjusted odds ratios (OR)s provided by logistic regression model using the interaction covariate between vaccination and study period

No.	症状	Study period: 0~6年				Study period: 3~6年				Study period: 3~9年				Study period: 0~9年*			
		95%信頼区間				95%信頼区間				95%信頼区間				95%信頼区間			
		OR	下限	上限	P値	OR	下限	上限	P値	OR	下限	上限	P値	OR	下限	上限	P値
1	月経不順	0.48	0.36	0.63	0.28	0.41	0.30	0.57	0.08	0.47	0.35	0.62	0.26	0.54	0.43	0.68	0.90
2	月経量の異常	0.78	0.47	1.27	0.77	0.63	0.36	1.10	0.33	0.61	0.38	0.98	0.17	0.76	0.51	1.12	0.56
3	関節やからだがかたくなる	0.74	0.47	1.16	0.64	0.57	0.33	0.98	0.63	0.69	0.44	1.10	0.73	0.89	0.62	1.28	0.07
4	ひどく頭が痛い	0.65	0.43	0.97	0.90	0.48	0.30	0.78	0.21	0.45	0.30	0.68	0.06	0.61	0.44	0.84	0.54
5	身体がだるい	1.37	0.94	2.00	0.00	1.23	0.79	1.92	0.01	1.14	0.78	1.66	0.01	1.27	0.94	1.70	0.00
6	すぐ疲れる	1.20	0.82	1.76	0.02	1.26	0.81	1.95	0.03	1.17	0.81	1.70	0.02	1.12	0.83	1.51	0.01
7	集中できない	1.04	0.66	1.62	0.03	1.30	0.77	2.20	0.01	1.02	0.65	1.60	0.05	0.81	0.56	1.16	0.22
8	視野の異常(暗くなる・狭くなるなど)	0.80	0.32	1.97	0.79	0.67	0.23	1.88	0.57	1.12	0.47	2.62	0.50	1.34	0.67	2.65	0.17
9	光を異常にまぶしく感じる	0.79	0.42	1.49	0.64	0.76	0.37	1.55	0.59	1.03	0.57	1.86	0.60	1.08	0.66	1.74	0.40
10	視力が急に低下した	0.23	0.14	0.36	0.01	0.24	0.14	0.43	0.04	0.37	0.22	0.61	0.55	0.34	0.23	0.51	0.29
11	めまいがする	0.94	0.65	1.36	0.07	0.89	0.58	1.37	0.19	0.92	0.63	1.33	0.07	0.98	0.73	1.30	0.01
12	足が冷たい	0.69	0.47	1.00	0.53	0.80	0.51	1.23	0.25	0.81	0.56	1.18	0.13	0.70	0.52	0.95	0.33
13	なかなか眠れない	0.57	0.36	0.92	0.12	0.52	0.30	0.90	0.09	0.95	0.60	1.48	0.39	1.04	0.72	1.50	0.11
14	異常に長く寝てしまう	0.83	0.57	1.20	0.63	0.93	0.60	1.43	0.37	1.04	0.72	1.49	0.07	0.93	0.69	1.24	0.14
15	皮膚が荒れてきた(湿疹・イボなど)	0.49	0.33	0.72	0.49	0.47	0.30	0.75	0.47	0.51	0.34	0.77	0.70	0.53	0.39	0.73	0.77
16	過呼吸	0.70	0.35	1.38	0.93	0.65	0.30	1.42	0.81	0.66	0.34	1.27	0.81	0.70	0.40	1.21	0.96
17	物覚えが悪くなった	3.92	1.82	8.50	0.00	2.73	1.14	6.54	0.07	2.50	1.25	4.98	0.04	3.59	2.05	6.25	0.00
18	簡単な計算ができなくなった	6.15	1.59	24.34	0.01	3.22	0.68	15.39	0.17	2.29	0.66	7.84	0.24	4.37	1.63	11.61	0.00
19	簡単な漢字が思い出せなくなった	1.89	0.77	4.60	0.22	1.16	0.42	3.23	0.85	1.73	0.76	3.94	0.16	2.82	1.46	5.39	0.00
20	身体が自分の意思に反して動く	2.97	0.72	12.32	0.38	2.99	0.63	14.30	0.42	3.21	0.97	10.62	0.21	3.19	1.17	8.66	0.12
21	普通に歩けなくなった	0.74	0.04	11.29	0.44	0.61	0.03	11.32	0.40	3.02	0.40	22.90	0.55	3.68	0.66	20.66	0.34
22	杖や車いすが必要になった	0.14	0.00	4.26	0.27	0.05	0.00	2.28	0.13	0.27	0.01	5.48	0.43	0.77	0.06	8.85	0.91
23	突然力が抜ける	1.45	0.46	4.52	0.91	1.87	0.52	6.64	0.65	4.05	1.53	10.73	0.01	3.16	1.42	6.96	0.01
24	手や足に力が入らない	1.64	0.63	4.28	0.58	2.45	0.84	7.16	0.23	2.74	1.17	6.41	0.05	1.83	0.90	3.70	0.22

OR: Odds ratio (オッズ比), P値: 接種とStudy periodの交互作用の検定, *: 八重・椿論文(文献3, open access) 中のTable 4より転載(licensed under CC BY-NC-ND 4.0)

としてStandardization法による年齢標準化リスク比の値を補足解析の位置づけで提示している。

以上のように、名古屋調査データにおいては交互作用(特に質的交互作用)の存在が想定される以上、年齢を共変量としたロジスティック回帰モデルによる解析結果の妥当性には問題があり、鈴木・細野論文がそのような問題のある解析結果を主解析結果としたうえで、“HPVワクチン接種と症状発現との間に関連がないことが示唆された”との結論を導いていることは不適切であると言わ

ざるを得ないのではないだろうか。なお鈴木・細野論文での結論の導き方に関する問題として、統計学的有意でないことをもって関連性がないと結論付けているという根本的問題も指摘しておきたい。2.1に紹介したとおり、鈴木・細野論文では、24症状の発現をprimary outcomeとして定義したうえで、ロジスティック回帰分析を用いた接種の影響(症状発現)に関する年齢調整オッズ比を24症状それぞれについて提示し、24症状全てにおいて統計学的有意に1を超えなかった(オッズ比

は、5症状では1以上で95%信頼区間は1をまたいでいる、19症状ではオッズ比が1未満で95%信頼区間は1をまたいでいない」という結果を提示している。そして結論には、“HPVワクチン接種と24症状発現との間で統計学的有意な関連は認められず、このことから、ワクチンと報告されている有害事象との関連はないことが示唆された。(翻訳、筆者)”と記述されており、これは明らかな統計的検定の誤用と考えられる。

さて、鈴木レターでは交互作用の検討の必要性を否定しているととれる主張がなされていたが、鈴木氏は朝日論座の中で、“鈴木論文では接種と接種年との交互作用を考え、この接種年で層別化したサブグループごとのオッズ比（非接種群に対する接種群の症状発現のオッズ比）を比較した”と主張している。そして“結果としてサブグループごとに異なるオッズ比が得られたものの、そのオッズ比の振れ幅は大きくなく、オッズ比が有意に2.0を越えた症状は1つだけであった”としている。この記述は、鈴木・細野論文でのサブグループ解析（コホート1～5での解析。コホートの抽出条件は2.1項を参照）での解析結果のコホート間での比較のことを指している。そしてこのコホート2～5での解析は、初回接種時年齢が13-14歳である集団の中において、接種年が2010年～2013年と変化することにより接種の影響がどのように異なるかをみたものとなっている。確かに、このサブグループ解析結果では、接種年が後になるほどオッズ比が大きくなる傾向が認められ、コホート2及び3（2010, 2011年に接種した集団）ではほとんどの症状でオッズ比が1を下回っていたのに対し、コホート4及び5（2012, 2013年に接種した集団）ではオッズ比が1を上回る症状の数が多くなっている。鈴木氏がこの結果をもって、接種による影響と接種年との交互作用を検討し、量的な交互作用の存在が想定されたと主張することは可能であろう。そして、この解析を量的な交互作用を考慮した解析として見るならば、その解析結果において、統計的有意に1を超えるオッズ比（コホート4での「突然力が抜ける」

のオッズ比1.6、コホート5での「物覚えが悪くなった」のオッズ比2.1）が得られている点は安全性シグナル検出の観点から非常に重要な知見ではないだろうか。

なお鈴木・細野論文では、このサブグループ解析について、接種と接種年との交互作用の検討のために実施したとは言及されていない。また鈴木・細野論文の考察では、接種年が後になるほどオッズ比が上昇した理由として、“接種年の違いはワクチン接種率の違いを反映しており、その結果として接種年が後の方でオッズ比が高めに見積もられた可能性がある”という解釈（筆者註：接種年が後の方の集団は接種率が低い集団であり、このような集団では非接種者の健康状態が相対的に良いことから、接種による症状発現のオッズ比が高く見積もられたという解釈と思われる）を提示しているのみである。つまりオッズ比が1を上回った理由として、いわゆる健康者接種バイアスの影響について言及しているものの、接種による危険性のシグナル検出の可能性については言及されていない。

また、鈴木・細野論文のサブグループ解析におけるコホート2～5は、調査時年齢でいうと18-19歳～15-16歳に対応しており、接種年の違いはすなわち調査時年齢の違いと捉えることが可能である。要するにコホート2～5のサブグループ解析とは、接種時年齢が一定（初回接種時年齢が13-14歳と一定）の集団の中で、調査時年齢（すなわち年齢）の違いでの層別解析を行ったと捉えることができる。ということは、コホート2～5を設定したサブグループ解析について鈴木氏は、“接種と接種年との交互作用を考えて行った”と主張しているが、要するにこれは接種と年齢との量的交互作用を検討していると言い換えることが可能である。このことは、鈴木氏がロジスティック回帰モデルでの主解析において接種と年齢との交互作用を検討しなかったこと、その必要性を否定していることと明らかに矛盾していると言えるのではないだろうか。

鈴木氏がサブグループ解析において接種と接種

年（調査時年齢と言ひ換えることが可能）との交互作用を検討したと主張するのであれば、交互作用の検定結果や量的、質的交互作用についての検討結果を提示しなかったのはなぜか、また主解析として行ったロジスティック回帰モデルでの解析で接種と年齢との交互作用の検討をせず、かつその検討の必要性も否定する理由は何か、鈴木氏からの合理的説明が求められるのではないだろうか。

この交互作用の問題に関しては、第24回日本薬剤疫学会学術総会での設楽・森川によるポスター発表¹⁷⁾においても指摘されている。設楽らは、名古屋調査データでは量的・質的交互作用が観察されたこと、特に質的な交互作用は効果の方向を逆転させるので注意すべきことなどを指摘している。また設楽らは、全年齢層を解析対象としたときに質的交互作用が認められたいくつかの症状（「物覚えが悪くなった」のGail-Simon検定でのP値0.002, 「突然力が抜ける」のGail-Simon検定でのP値0.048）では、健康者接種バイアスの影響が少ないと考えられる低年齢層での解析で、接種の影響をうかがわせる結果が得られたことも報告している。

最後に、非接種群に対する接種群の症状発現オッズ比が統計的有意に1を超えた症状として、八重・椿解析では「物覚えが悪くなった」、設楽・森川によるポスター発表においては「物覚えが悪くなった」と「突然力が抜ける」という症状が挙げられており、さらに鈴木・細野論文のサブグループ解析では、オッズ比が統計的有意に1を超えた症状として「物覚えが悪くなった」と「突然力が抜ける」が示されている。すなわち八重・椿論文、設楽・森川ポスター、及び鈴木・細野論文において、接種による危険性のシグナルが検出された症状には共通性があることがわかる。そしてこれらの症状がHPVワクチン接種後症状に特徴的の症状として報告されているものであることを考慮すれば、これらの知見には危険性シグナルの検出の点から重要な意味があると考えるのが妥当ではないだろうか。

3.4 年齢層別解析における対照群の設定

鈴木レターでは、“全年齢層の接種群と15-16歳年齢層の非接種群との比較は不適切な比較であり、15-16歳年齢層の非接種群を対照群とする場合には15-16歳の接種群のみを使用するべきである”と指摘している。

2.1項に記述したとおり、八重・椿論文では、年齢層別での解析結果を提示しており、鈴木氏が指摘する、15-16歳年齢層での接種群と非接種群の比較結果も提示している。そして、全年齢層の接種群と15-16歳年齢層の非接種群との比較は、健康者接種バイアスの影響を考慮した解析の手法として行ったものであり、探索的解析手法の一つとして妥当であると考えられる。

名古屋調査の15-16歳の非接種群は、健康状態不良のために接種できなかったというよりも、健康でかつ接種しなかった人という想定が可能と考えられた。このような想定が可能となる理由としては、1つには、この年齢層では非接種者割合が18歳以上の集団（非接種者割合は約10～15%）よりもはるかに高かった（15歳で約85%、16歳は約50%）点が挙げられる。つまり非接種者割合が比較的高いことは、15-16歳集団では健康上の理由で接種できなかった人だけでなく、健康でありながらも接種しなかった人が一定割合存在したことを意味しており、したがって15-16歳の非接種群は健康でかつ接種しなかった人（が比較的多い）と想定できることになる。また、この15-16歳の集団が定期接種対象年齢に達した時期には副反応報道や厚生労働省による接種の積極的勧奨差し控え通知があったことから判断すると、自主的判断により接種を控えた人が多かった（健康上の理由で接種できなかった人は比較的小数であった）のではないかと考えられ、この点も、理由の一つとして挙げられるであろう。すなわち15-16歳の非接種群は、健康でかつ接種しなかった人（が比較的多い）と想定することは妥当であると考えられ、健康で接種が可能であった全年齢層接種群に対し、15-16歳の非接種群を対照として設定することにより、健康者接種バイアスの影響が比較的

少ない状態での比較が可能となるのであり、このような比較での解析は探索的解析として許容されると考える。

なお健康者接種バイアスへの対処方法としては、たとえば傾向スコアを応用することも考えられる。健康状態等が影響すると想定される接種（の受けやすさ）の確率を傾向スコアとして推定し、このスコア値で調整することで健康者接種バイアスの影響を考慮する解析方法である。名古屋調査データには対象者の基礎疾患などの情報が無いことから、健康状態を変数として用いて算出する傾向スコアを求めることは実施困難と考えられた。ただし、接種に影響する要因としては個人の健康状態だけでなく、たとえばHPVワクチン接種関連の各種政策の実施時期とその当時における対象者の年齢との関係に着目して得られる変数等を用いて傾向スコアを算出するという可能性は今後の検討課題と考える。

3.5 検定の多重性の調整

鈴木レターでは、24症状について検定を行う場合は検定の多重性の問題が生じるため多重性の調整が必要であるとの主張がなされ、この点は、その後の朝日論座や雑誌小児科でも繰り返し鈴木氏により指摘されている。八重・椿論文では、方法の統計解析の項で多重性の調整を行わない旨を述べたうえで、結果には多重性の調整を行わない値としてのオッズ比とその95%信頼区間を提示した。複数の検定を実施した場合に検定の多重性が存在することは鈴木氏の指摘のとおりである。しかし筆者は、多重性の調整が求められるのは明確な研究仮説設定のもとに介入の有効性を検証するような検証的試験の場合であって、探索的な検討を主目的とするような観察研究、特に安全性問題に関する探索的分析を行う研究においては、多重性の調整は必要でなく、多重性の調整を行わない解析結果の提示のほうが適切であると考え、この根拠については4の項で詳述する。

なお鈴木・細野論文では、方法論の項に検定の多重性の調整についての言及はなされておらず、

結果においても多重性の調整を行ったという説明はないため、多重性の調整は実施されていないと思われる。2.1項に紹介したとおり、鈴木・細野論文では、24症状について症状発現を主要転帰とした場合のロジスティック回帰モデルによる年齢調整オッズ比を主解析結果（オッズ比とその95%信頼区間）として提示し、24症状中5症状ではオッズ比が1以上であるが95%信頼区間が1をまたいでおり統計学的有意ではなかったこと、残り19症状ではオッズ比が1未満であり、うち14症状では統計学的有意、5症状では有意ではなかったことを報告している。また病院受診をアウトカムとした場合の年齢調整オッズ比（及びその95%信頼区間）の結果では、24症状のうち「月経量の異常」、「月経不順」、「ひどく頭が痛い」の3症状でオッズ比が1を上回り統計学的有意であったことも報告されている。すなわち鈴木・細野論文は、検定の多重性の調整については言及しないまま（おそらく多重性の調整は行われず）、結果をオッズ比とその95%信頼区間で提示し、統計的検定の考え方に基づいて統計的有意か有意でなかったかの結果を示している。鈴木・細野論文では検定の多重性をどのように考え、どのような対処を行ったかが不明である。これらの点について説明が求められるであろう。

3.6 利益相反の開示

鈴木氏は、“八重・椿論文の筆頭著者である八重は薬害オンブズパースン会議のメンバーであったにもかかわらず、論文中にこの事実が明記されていないことが問題である”と鈴木レターにおいて指摘している。

筆者は、八重・椿論文でのCOI開示については、その手続きに不備はないと考える。八重は論文投稿時に、当時薬害オンブズパースン会議のメンバーであったことを申告しており、この事実が論文中に記述されなかった理由は、雑誌JJNSの規定（金銭的關係がある場合には論文中に組織と関係があることを記述する）にのっとった結果である。このことは編集長からのレターでも明確に説

明されている⁶⁾。投稿時における著者と一組織との関係性については、金銭的關係がなくとも論文中に明記すべきかどうかについては議論のあるところと考える。しかし今回の八重・椿論文の場合は、雑誌側が規定にのっとり、論文中への記述を求めなかったということが事実であり、八重のCOI申告に関する姿勢と行った手続きには全く問題は無いことを再確認しておきたい。

4. 安全性の分析において、検定の多重性を調整することの問題点

以上、2019年1月に八重・椿論文が公表されてから今日までの間に行われた鈴木氏による八重・椿論文への指摘の主なもの、及びそれらを踏まえての八重・椿論文著者としての筆者の見解について詳述した。ここでは、鈴木・細野論文、八重・椿論文の両方に関連する検定の多重性の調整に関しては特に重要な問題であると考え、安全性の分析、特に危険性のシグナル検出の観点からその問題点についての解説を試みたい。

筆者はまず、有効性評価と安全性評価では、統計的検定を用いる場合の考え方が異なることが認識される必要があると考える。すなわち有効性評価の場合は検証の立場で検定の多重性の調整を行うことが求められるが、安全性評価の場合は探索的立場で検定を行い、必ずしも多重性の調整を求めない姿勢が重要であると考え。このような考え方に立たなければ、消費者危険（有効ではない医薬品を医療消費者が使うことになる、あるいは安全性が十分担保されていない医薬品を消費者が使うことになる）を一定の確率以下に保証することができないことは、椿により以前から指摘されてきたことである¹⁸⁾。たとえば、仮に有効性評価において検定の多重性の調整を行わないとすると、臨床試験の中でたくさんのエンドポイントでの検定を行い（多項目検定による検定の多重性が存在することになる）、どれかのエンドポイントで有意差が認められれば有効な医薬品であるという主張が可能となり、結果として甘い基準での有

効性検証を可能としてしまうことになる。一方、安全性評価において検定の多重性の調整を行うとすると、臨床試験の中で複数の危険性のシグナルが出ていたとしても、多重性の調整の結果としてより厳しい基準で検定することとなり、結果としてシグナルを見逃すことにつながってしまう。すなわち多重性の調整の適用を誤ると（有効性評価で多重性の調整をしない、安全性評価で多重性の調整を行う）、患者に有効性が不十分な医薬品が提供される、あるいは患者が安全性の十分担保されていない医薬品を使うことになる、このような危険性の確率を一定以下に抑えることができなくなることを意味するのである¹⁸⁾。

なお同様の考え方は、欧州医薬品庁の医薬品委員会（Committee for Proprietary Medicinal Products：CPMP、2002年当時。現在はCommittee for Medicinal Products for Human Use：CHMP）が2002年に出した「臨床試験での多重性問題における注意点（Points to Consider on Multiplicity Issues in Clinical Trials）」でも明確に述べられている。「2.4 安全性に関する変数における多重性（Multiplicity in safety variables）」には、「In those cases where a large number of statistical test procedures is used to serve as a flagging device to signal a potential risk caused by the investigational drug it can generally be stated that an adjustment for multiplicity is counterproductive for considerations of safety.」と記述され、多重性の調整を行うことは安全性シグナル検出を阻害する（検出においては逆効果となる）ことが指摘されている¹⁹⁾。また近年では、*New England Journal of Medicine*の統計報告ガイドライン（2019年改定版）においても、臨床試験の項で「If significance tests of safety outcomes (when not primary outcomes) are reported along with the treatment-specific estimates, no adjustment for multiplicity is necessary.」と記述され、安全性アウトカムの検定においては多重性の調整は必要ないことが明記されている²⁰⁾。

5. 最後に

筆者はHPVワクチンに関しては、現在のワクチンの承認自体を否定するものではなく、副反応の高リスク集団を特定して接種対象から除外することで、より安全な接種となる方策が求められると考える。すなわち有効性に関する知見を臨床試験結果から正確に把握するとともに、安全性に関する探索的研究により、重篤な副反応が出る対象者層を明らかにしていくこと、そして重篤な副反応のリスクが高い対象者を接種対象者から除くことができるような対策を検討すること、これらが接種を進めるうえでの大前提として求められると考える。八重・椿論文では、名古屋調査データを用い、年齢データやStudy periodで分類した層での接種による接種後症状発生への影響に着目し、HPVワクチンは安全である（ワクチン接種と症状経験との間に関連はない）という仮説は棄却されるのか（危険性のシグナルが検出されるのか）という観点から統計的に探索した。すなわち八重・椿論文は、安全性に関する探索を目的とした探索的解析研究の位置づけで実施したものであり、妥当な疫学、統計学的手法にのっとったものであると考える。ただし八重・椿論文の結果においてシグナルとして観察されたHPVワクチンの危険性については、さらなる検討が必要であることは言うまでもない。

一方、鈴木・細野論文は、24症状ごとの症状発生を主要転帰とし、ロジスティック回帰モデルを用いてHPVワクチン接種による各症状発生の年齢調整オッズ比を求め、これを主解析結果として提示したうえで、症状発生のオッズ比が統計的に有意に1より高い値を示した症状はなかったと報告し、結論では、ワクチンと報告されている有害事象との関連はないことが示唆されたとしている。また考察においては、得られた結果は日本におけるHPVワクチンの安全性を示唆していると言及している。さらに標題には“*No association*”との記載もある。統計的検定結果において統計的

に有意な差が認められないことが同等であることを意味するわけではなく、安全性評価における統計的有意差なしの結果は、介入群と対照群の間で安全性が同等であったことを意味しない。しかしながら鈴木・細野論文では、統計的有意差が認められなかったという結果から、接種と症状発生との関連はなく安全性が示唆されるとの結論を導いており、これは統計的方法の誤用に基づいた不適切な主張であると考えられる。

なお、名古屋市が公開したデータの中には調査回答者による自由記載欄の記述も含まれているが、個々の記述内容が接種者によるものか非接種者による回答かの区別はできるものの、それ以外の対象者情報が紐づけされていないため、接種後症状と思われる記述に関する具体的な分析ができないデータとなっていることは非常に惜まれる点である。

HPVワクチンが子宮頸がん予防のために適切に使用されるためには、有効性の検証はもちろんのことであるが、安全性の確認をおろそかにすることがあってはならない。名古屋市が約7万1千人の若年女性を対象として実施した大規模調査から得られた貴重なデータを無駄にすることなく、できる限り適切な統計解析手法を適用して安全性に関する知見を得る努力をすることが疫学、統計学の研究者には求められているのではないだろうか。本稿が名古屋調査データに関する健全な科学的議論の基礎となり、HPVワクチンの適正使用に資するものとなることを期待するものである。

利益相反

本稿の内容と関連して筆者に開示すべき利益相反はない。

文 献

- 1) 健康福祉局感染症対策室感染症係, 名古屋市. 子宮頸がん予防接種調査結果 [cited 2021 August 17]. Available from: <https://www.city.nagoya.jp/kenkofukushi/page/0000088972.html>

- 2) Suzuki S, Hosono A. No association between HPV vaccine and reported post-vaccination symptoms in Japanese young women: Results of the Nagoya study. *Papillomavirus Research*. 2018; 5, 96-103. [cited 2021 August 17]. Available from: <https://doi.org/10.1016/j.pvr.2018.02.002>
- 3) Yaju Y, Tsubaki H. Safety concerns with human papilloma virus immunization in Japan: Analysis and evaluation of Nagoya City's surveillance data for adverse events. *Japan Journal of Nursing Science*. 2019 Jan 28 [cited 2021 August 17]. Available from: <https://doi.org/10.1111/jjns.12252>, licensed under CC BY-NC-ND 4.0.
- 4) Suzuki S. Letter to the editor: Safety concerns with human papilloma virus immunization in Japan: Analysis and evaluation of Nagoya City's surveillance data for adverse events. *Japan Journal of Nursing Science*. 2019 Aug 26 [cited 2021 August 17]. Available from: <https://doi.org/10.1111/jjns.12258>
- 5) Yaju Y, Tsubaki H. Letter to the editor: Authors' response to letter to the editor: Safety concerns with human papilloma virus immunization in Japan: Analysis and evaluation of Nagoya City's surveillance data for adverse events. *Japan Journal of Nursing Science*. 2019 Aug 26 [cited 2021 August 17]. Available from: <https://doi.org/10.1111/jjns.12275>
- 6) Holzemer W. Editor's Reply to Letter to the Editor of Dr. Suzuki. *Japan Journal of Nursing Science*. 2019 Aug 26 [cited 2021 August 17]. Available from: <https://doi.org/10.1111/jjns.12282>
- 7) 鈴木貞夫. HPVワクチンと名古屋スタディ／上 [cited 2021 August 17]. Available from: <https://webronza.asahi.com/national/articles/2020012100004.html>
- 8) 鈴木貞夫. HPVワクチンと名古屋スタディ／下 [cited 2021 August 17]. Available from: <https://webronza.asahi.com/national/articles/2020012100012.html>
- 9) 鈴木貞夫. 名古屋スタディの概要と研究をめぐる状況・展望. *小児科*. 2021 ; 62(6) : 562-7.
- 10) 佐藤俊哉. 交絡 事実と反事実の比較. *科学*. 2008 ; 78(4) : 427-9.
- 11) 木原正博, 木原雅子. 1.2 年齢効果, コホート効果, 歴時効果の分析. In : 木原正博, 木原雅子, 訳. アドバンスト分析疫学 369の図表で読み解く疫学的推論の論理と数理. 東京 : 株式会社メディカル・サイエンス・インターナショナル ; 2020. P. 3.
- 12) 山口拓洋. 脳卒中研究に必要な統計学のABC ロジスティック回帰分析. *分子脳血管病*. 2014 ; 13(1) : 93-6.
- 13) 丹後俊郎. 7.2.2 交互作用は調整できる?. In : 丹後俊郎, 著. 統計学のセンス デザインする視点・データを見る目. 東京 : 朝倉書店 ; 1998. P. 116-7.
- 14) 佐藤俊哉. 疫学研究における交絡と効果の修飾. *統計数理*. 1994 ; 42(1) : 83-101.
- 15) Miettinen, O S. Confounding and effect-modification. *American Journal of Epidemiology*. 1974; 100: 350-3.
- 16) Greenland, S. Modeling and variable selection in epidemiologic analysis. *American Journal of Public Health*. 1989; 79: 340-9.
- 17) 設楽 敏, 森川敏彦. 「名古屋市子宮頸がん予防接種調査」データに潜むバイアスを探る. 日本薬剤疫学会第24回学術総会ポスター発表2018, P. 1-7.
- 18) 椿 広計. 5.2推論の多重性による消費者危険の増大. In : 椿 広計, 藤田利治, 佐藤俊哉, 編. これからの臨床試験 医薬品の科学的評価-原理と方法. 東京 : 朝倉書店 ; 2000. P. 69-72.
- 19) COMMITTEE FOR PROPRIETARY MEDICINAL PRODUCTS (CPMP). POINTS TO CONSIDER ON MULTIPLICITY ISSUES IN CLINICAL TRIALS ; 19 Sep 2002 [cited 2021 August 17]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-multiplicity-issues-clinical-trials_en.pdf
- 20) Statistical Reporting Guidelines, *New England Journal of Medicine* [cited 2021 August 17]. Available from: <https://www.nejm.org/author-center/new-manuscripts>
- 21) 寺尾 哲, 森川敏彦. 15.5 交互作用. In : 寺尾 哲, 森川敏彦, 共著. 生物統計学 標準教科書. 東京 : ムイスリ出版株式会社 ; 2018. P. 251-5.
- 22) 寒水孝司. 臨床試験における多重性の諸問題. *計量生物学*. 2015 ; 36(Special Issue) : S 87-98.

(投稿日 : 2021年 8月 17日)

(受理日 : 2021年 11月 6日)

(誤記訂正 : 2022年 3月 9日)