

Distortion in Risk Estimation Driven by Study Period Adjustment and Young-Age Event Sparsity: A Critical Evaluation of Secondary Analyses on the Nagoya City HPV Vaccination Survey

Sadao Suzuki

ssuzuki@med.nagoya-cu.ac.jp

Nagoya City University

Research Article

Keywords: Information bias, Epidemiological methods, Odds ratio distortion, HPV vaccination, Study period adjustment, Secondary data analysis

Posted Date: October 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11007157/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

- 1 **Distortion in Risk Estimation Driven by Study Period Adjustment and**
- 2 **Young-Age Event Sparsity: A Critical Evaluation of Secondary**
- 3 **Analyses on the Nagoya City HPV Vaccination Survey**
- 4
- 5 Sadao Suzuki, M.D., M.Sc., Ph.D.
- 6 Professor Emeritus, Nagoya City University
- 7 Former Professor, Department of Public Health, Nagoya City University Graduate
- 8 School of Medical Sciences
- 9 Email Address: wisesada@hotmail.com, ssuzuki@med.nagoya-cu.ac.jp
- 10 ORCID : 0000-0001-5988-4849

Abstract

Background

This study evaluates the methodological validity and mathematical properties of the Study Period (SP) adjustment method introduced in a high-profile epidemiological analysis by Yaju and Tsubaki, which challenged the primary findings of the Nagoya Study on HPV vaccine safety. While independent experts have raised concerns regarding systematic information bias within this covariate, its precise mathematical mechanism and structural impacts on risk estimation remain fully unelucidated.

Methods

We conducted a rigorous cross-model comparison across the three analytical frameworks (Models 1–3) and findings (Table 4) presented by Yaju and Tsubaki. Using the original baseline benchmark from the Nagoya Study primary analysis (OR_{NS}), we calculated relative odds ratios and utilized Pearson correlation coefficients to assess how event deletion and adjustment-factor alterations systematically interact with pre-vaccination symptom baseline proportions (EGP%).

Results

Methodological flaws escalated sequentially through each analytical layer. Model 1 systematically underestimated risks for high-EGP% symptoms through asymmetric event deletion within the vaccinated cohort ($r = -0.96$ with EGP%). Conversely, Model 2 introduced SP adjustment instead of standard age adjustment, creating a structural asymmetry in observation windows that artifactually diluted onset density during youth and inflated risks exclusively for young-age event-sparse symptoms. Model 3 incorporated an unjustified interaction term, forcing a fictitious negative gradient over time ($r = 0.78$) that mathematically guaranteed maximum extreme estimates at $SP = 0$, presented as representative risks in Table 4. Consequently, the correlation with the valid benchmark OR_{NS} dropped precipitously from 0.84 (Model 1) to 0.38 (Model 2), and ultimately to 0.27 (Table 4). This structural bias produced severe, bidirectional distortions: the young-age event-sparse Symptom 18 was artificially inflated from a significant lower risk of 0.70 (OR_{NS}) to 4.37, while the high-EGP% Symptom 10 was deflated to 0.34. The authors selectively reported these mathematical artifacts while ignoring clinically contradictory results.

Conclusions

47 The secondary analytical models and Table 4 estimates presented by Yaju and Tsubaki
48 are artifacts driven by structural information bias and invalid mathematical
49 extrapolation. Their conclusions must be formally re-evaluated to preserve scientific
50 integrity.

51

52 **Key words**

53 Information bias, Epidemiological methods, Odds ratio distortion, HPV vaccination,
54 Study period adjustment, Secondary data analysis

1. Introduction

This paper provides a rigorous mathematical and epidemiological evaluation comparing the primary analysis of the 2015 Nagoya City HPV Vaccination Symptom Survey (hereafter, the “Nagoya Study”) with three alternative secondary statistical models subsequently proposed by Yaju and Tsubaki. Our primary objective is to interrogate the methodological rigor of these competing analytical approaches, focusing critically on whether the calculated odds ratios for the 24 tracked symptoms serve as valid indicators of causal relationships or vaccine-associated risk alterations. Specifically, we demonstrate the precise mathematical mechanisms through which the interaction between the temporal distribution of symptom onset—characterized by young-age event sparsity—and specific statistical constraints, particularly the “Study Period” (SP) adjustment, artifactually distorts the analytical findings.

2. Methods and Definitions of Key Indicators

In this section, we define the mathematical indicators used to compare and validate the primary analysis of the Nagoya Study (hereafter, the “Nagoya primary analysis”) and the three models proposed in the Yaju and Tsubaki paper [2]. The specific numerical data utilized for this evaluation were extracted from the published findings of both

papers and are comprehensively summarized in Table 1.

(1) Notation of Odds Ratios for Each Model

Throughout this paper, the odds ratios (ORs) calculated for the 24 symptoms across the respective models are abbreviated and denoted as follows:

- OR_{NS} : Odds ratio from the Nagoya primary analysis (Table 3).
- OR_{M1} , OR_{M2} , OR_{M3} : Odds ratios from Model 1 (Table 2), Model 2 (Table 3), and Model 3 (Table 4) of the Yaju and Tsubaki paper, respectively.

(2) Definition of Grace Period (GP) and Event Rate During Grace Period (EGP%)

A core methodological characteristic of the analysis by Yaju and Tsubaki is the exclusion of symptom onsets occurring prior to the initial vaccination within the vaccinated cohort, which they operationalized as “symptoms unrelated to causality”.

Although the Nagoya survey data recorded birth years, the current study standardizes age by the age attained within each academic year; for instance, sixth-grade elementary school students are designated as 12 years old. Accordingly, we define the interval from

the baseline symptom onset age (age 12) until the receipt of the initial vaccination as the “Grace Period” GP. Symptoms occurring during this interval are denoted as “Events during Grace Period” EGP. The proportion of EGP relative to the total symptomatic cases for each symptom EGP% is calculated using the following formula:

$$EGP\% = \frac{N_{Nagoya} - N_{YT}}{N_{Total}}$$

Where:

- N_{Nagoya} is the number of vaccinated symptomatic cases in the Nagoya Study,
- N_{YT} is the number of vaccinated symptomatic cases in the Yaju and Tsubaki paper, and
- N_{Total} is the total number of symptomatic cases in the Nagoya Study.

For instance, for Symptom 1, the calculation yields $EGP\% = (5,468 - 3,603) / 5,912 = 31.6\%$. This EGP% serves as an index tracking the natural background incidence of symptoms in early adolescence, reflecting a tendency to manifest during this growth phase independent of vaccination status. Conversely, a low EGP% combined with infrequent symptom onset during the GP is operationalized as “young-age sparsity.” In

Table 1, the 24 symptoms are sorted in ascending order of EGP%, which inversely reflects the degree of young-age sparsity.

(3) Four “Relative Odds Ratios” for Evaluating Model Concordance and Intra-Model Changes

To quantify how variations in model specifications and variable selection across the mathematical models affect the calculated odds ratios, we introduce four relative indicators. The geometric relationships between these indicators and EGP% are illustrated via scatter plots in Figure 1. Among the 24 symptoms evaluated, the five symptoms that exhibited statistically significant elevations in odds ratios in the final model (Table 4) of the Yaju and Tsubaki paper are denoted by closed circles (●), whereas the four symptoms showing significantly decreased odds ratios are denoted by open circles (○). The remaining 15 symptoms, which showed no statistically significant changes, are denoted by small gray circles (●). These nine key symptoms are highlighted in bold in Table 1. To account for the mathematical asymmetry of odds ratios, log-transformed values of these relative indicators were utilized to compute the Pearson correlation coefficients.

1. **OR_{M1}/OR_{NS} Ratio (Relative Odds Ratio of Model 1 to the Nagoya Primary**

Analysis): Because both models account for age distribution, this ratio dynamically approaches 1.0 in the absence of EGP. The primary utility of this indicator is to evaluate the magnitude of the mathematical compression in odds ratios for symptoms characterized by a high EGP%.

2. **OR_{M2}/OR_{M1} Ratio (Relative Odds Ratio of Model 2 to Model 1):** Since

both models share an identical procedure for EGP exclusion, calculating this ratio effectively cancels out the mathematical effect of EGP removal.

Consequently, it serves as a pure and direct measure of the divergence caused solely by the substitution of adjustment factors (i.e., standard age-adjustment versus the controversial SP-adjustment).

3. **OR_{M3}/OR_{NS} Ratio (Relative Odds Ratio of Model 3 to the Nagoya Primary**

Analysis): This index quantifies the cumulative, overall divergence between the final models of the respective studies, reflecting the combined impact of all methodological modifications.

4. **IM_{M3} (Increment Multiplier of the Model 3 Odds Ratio):** Under the

assumptions of Model 3, the relative odds ratio per unit increase in SP is constrained to be constant. This indicator represents the exponentiated slope

(constant multiplier) of the regression line derived from log-odds, computed from the SP-specific odds ratios published in Tsubaki's commentary [3] (Table 1). For instance, for Symptom 18, the ratio of the odds ratio at SP=1 to that at SP=0 yields $3.20/4.37 = 0.73$, demonstrating that the odds ratio consistently decreases by a factor of 0.73 per unit increase in SP.

3. Methodological Validity of the Nagoya Study Primary Analysis and Age

Adjustment

This section reviews the methodological and mathematical rigor of the primary analysis of the Nagoya Study, which serves as the fundamental baseline for our evaluation.

Academic discussions surrounding this primary analysis typically involve two distinct issues: (1) the validity of age-adjusting the odds ratios, and (2) the appropriateness of presenting a single, age-adjusted summary measure in the presence of potential age-exposure interactions. The current study focuses primarily on the former. Regarding the latter, Subgroup Analysis 1 of the original Nagoya Study has already confirmed that vaccination does not pose an elevated risk within any age-specific cohort. This ensures that there is no interpretive discrepancy between the stratified cohorts and the primary analysis, thereby validating the practical and clinical appropriateness of presenting a

single pooled estimate.

The Nagoya Study intentionally omitted unadjusted univariate analysis. In analytical epidemiology, age adjustment is standard practice to eliminate confounding and isolate the true exposure effect. Because the primary objective of the survey was to quantify vaccine-associated risk rather than to evaluate the independent effect of age itself, presenting confounded crude results was unnecessary. Omitting unadjusted univariate outcomes and presenting only multivariable-adjusted findings is routine and epistemologically sound in analytical epidemiology, unless the investigation of the confounding mechanism itself is the core objective of the study.

Conversely, Yaju and Tsubaki argued in their discussion that “with respect to a confounder, the data did not meet the assumption.” However, if age were truly not a confounding factor, the substantial divergence between the unadjusted univariate results and the age-adjusted odds ratios presented in their own Table 2 would remain mathematically inexplicable, thereby rendering their assertion logically untenable.

Furthermore, the Pearson correlation coefficient between OR_{NS} and EGP% was 0.02,

demonstrating that the magnitude of EGP% exerted no mathematical influence on the primary analysis odds ratios. This empirical finding explicitly indicates that while the inclusion of causally unrelated pre-vaccination symptoms among vaccinated individuals was methodologically questioned by subsequent authors, the baseline proportion of these pre-exposure cases did not introduce any systematic bias into the primary analysis results.

Consequently, the Nagoya primary analysis contains no structural or methodological flaws that undermine its internal validity as a comprehensive assessment. Indeed, no peer-reviewed rebuttal letters or formal correspondences challenging the epidemiological validity of the original Nagoya Study have been published in the literature. In the subsequent sections, we therefore treat the Nagoya primary analysis as a valid and reliable benchmark against which to evaluate the concordance and divergence of the alternative models proposed by Yaju and Tsubaki.

4. Critical Evaluation of the Three Yaju & Tsubaki Models and Their Results

(1) Methodological Flaws Regarding Pre-Vaccination Symptoms (EGP) in Model 1

Model 1 of the Yaju and Tsubaki paper excludes EGP cases from the vaccinated cohort while keeping all other analytical structures identical to the Nagoya primary analysis. Consequently, the Pearson correlation coefficient between OR_{M1} and OR_{NS} remains notably high ($r = 0.84$). However, because this event deletion process is applied exclusively to the vaccinated cohort, it introduces a structural asymmetry that artifactually underestimates risks relative to OR_{NS} . The magnitude of this distortion depends directly on the baseline symptom frequency: for instance, while approximately 31.6% of symptomatic cases are deleted for Symptom 1 due to its high EGP%, the deletion rate remains under 10% for six specific symptoms (Symptoms 17, 18, 19, 21, 23, and 24). Crucially, the relative odds ratio (OR_{M1}/OR_{NS}) exhibits an extremely strong negative correlation with EGP% ($r = -0.96$). This mathematical relationship demonstrates that the calculated odds ratios decrease linearly as a direct function of the increasing baseline proportion of EGP, rendering Model 1 systematically biased against symptoms with higher pre-vaccination frequencies.

(2) Pre-Vaccination Symptoms and the Study Period Issue in Model 2

Model 2 introduces SP adjustment in place of standard age adjustment. Because both Model 1 and Model 2 share the identical premise of excluding EGP from the vaccinated

cohort, calculating the OR_{M2}/OR_{M1} ratio mathematically cancels out the effect of EGP exclusion, thereby isolating the direct impact driven solely by the alteration in adjustment factors (i.e., standard age versus SP). Strikingly, despite this cancellation, a strong negative correlation ($r = -0.68$) persists between EGP% and the OR_{M2}/OR_{M1} ratio. Specifically, for Symptom 1, which exhibits the highest EGP% (31.6%), the OR_{M2}/OR_{M1} ratio is exactly 1.00 ($OR_{M1} = 0.55$, $OR_{M2} = 0.55$), demonstrating zero alteration attributable to the choice of adjustment factor. Conversely, for Symptoms 18 and 21, where the EGP% is less than 9%, the ratios show substantial artifactual increases exceeding 1.85-fold. Furthermore, the correlation coefficient between OR_{M2} and EGP% is also strongly negative ($r = -0.68$), and its correlation with the baseline benchmark OR_{NS} drops sharply to 0.38, in stark contrast to the 0.84 observed for OR_{M1} .

The SP variable was originally introduced under the methodological premise that it would yield conceptually more valid odds ratios by controlling for causally unrelated pre-vaccination symptoms. However, the empirical behavior of the data directly invalidates this rationale. While high-EGP% symptoms—where EGP exclusion should theoretically exert the largest mathematical impact—exhibit minimal divergence from

the standard age-adjusted model, young-age event-sparse symptoms—where the impact of EGP removal should be entirely negligible—demonstrate a striking, disproportionate divergence. The root cause of this paradox lies in the profound information bias inherent to the operational definition of SP. The observation window for the vaccinated cohort is heavily restricted to the interval from the initial vaccination to the 2015 survey, whereas that for the unvaccinated cohort asymmetrically spans the entire period from age 12 to the 2015 survey. This structural asymmetry in the observation windows yields fundamentally disparate, non-comparable denominator definitions between the two comparison groups.

Figure 2 schematically illustrates the comparisons at identical ages (A versus B) and identical SPs (B versus C) for young-age event-sparse symptoms. Analytical comparisons initiating from the solid triangles (▼) originate uniformly at age 12, thereby rendering the baseline young-age event sparsity mathematically comparable between the comparison groups. Conversely, comparisons initiating from the open triangles (▽) introduce differing baseline onset ages into the unvaccinated group, causing the young-age event sparsity to manifest disproportionately within the unvaccinated cohort. This asymmetric young-age sparsity artifactually dilutes the

overall onset density in the unvaccinated group—independent of any actual vaccine effect—thereby generating a systematic bias that overestimates the resulting odds ratios. For symptoms characterized by a high EGP%, however, such dilution of the onset density does not occur within the unvaccinated group, resulting in the observed minimal divergence from the standard age-adjusted model. This empirical evidence definitively demonstrates that SP adjustment possesses a fundamental, irreconcilable contradiction: it produces empirical phenomena entirely opposite to its intended methodological purpose. Due to these methodological and mathematical breakdowns, the SP adjustment method in Model 2 is fundamentally invalid. Furthermore, regarding the information bias inherent to the SP variable, explicit cautionary notes directly evaluating the Yaju and Tsubaki paper have already been archived in the literature, recognizing it as a “rare empirical example where information bias exists within a covariate and distorts results.”

Regarding the supplementary analysis restricting the unvaccinated controls to ages 15–16 (Table 6), a strong negative correlation ($r = -0.73$; data not shown) was observed between EGP% and the relative odds ratios of Table 6 to the Nagoya primary analysis (the OR_{T6}/OR_{NS} ratio). Although Yaju and Tsubaki attributed the lower odds ratios in Table 6 to the “healthy vaccinee bias,” this assertion is highly speculative and entirely

lacks empirical backing. Rather, our findings demonstrate that it merely reflects the identical information bias driven by the interaction between the inherent young-age event sparsity of the examined symptoms and the non-comparable window constraints.

(3) Mechanism of Odds Ratio Distortion in Model 3

Model 3, the final model presented by Yaju and Tsubaki, assumes a constant rate of change in the odds ratios per unit increase in SP via an interaction term. Plotting this constant rate of change—operationalized as the increment multiplier (IM_{M3})—against EGP% for each of the 24 symptoms revealed a robust, statistically significant positive correlation ($r = 0.78$; Figure 1C). This strong correlation indicates that a lower EGP% (i.e., more pronounced young-age event sparsity) produces a steeper, downward-sloping negative gradient relative to SP unit changes. Consequently, this mathematical property forces the calculated odds ratios to reach their absolute maximum precisely at the minimum SP value of 0. This phenomenon is entirely explicated by the underlying geometric structure of the open data.

While the unadjusted SP for the unvaccinated cohort ranges from 3 to 9 years, the baseline observation window for all subjects remains anchored at the unified age of 12.

Figure 3 schematically illustrates how the mathematical impact of young-age event sparsity within the unvaccinated cohort varies dynamically across different SP values, utilizing $SP = 3, 6,$ and 9 as comparative baselines. For symptoms characterized by pronounced young-age event sparsity (such as Symptoms 17, 18, 19, and 23), individuals with larger SP values possess a relatively smaller proportion of the sparse young-age period (represented by the light-shaded areas in Figure 3) within their cumulative SP span. For instance, assuming a 1-year sparse onset period starting at age 12, its relative duration within the total observation window systematically decreases to $1/3, 1/6,$ and $1/9$ for $SP = 3, 6,$ and $9,$ respectively. This inversely proportional contraction of the sparse period relative to an increasing SP mathematically attenuates the artificial overestimation of the odds ratios.

Table 1 of the Tsubaki editorial presents the SP-specific odds ratios across 10 strata for all 24 symptoms under Model 3; of these, the trajectories for four representative symptoms are illustrated in Figure 4. As clearly demonstrated in the figure, symptoms with high EGP% (#1 and #10) exhibit virtually no variation across different SP values, resulting in an IM_{M3} close to 1. In contrast, symptoms with pronounced young-age

event sparsity (#17 and #18) display a distinct negative slope as SP increases, which corresponds to a smaller IM_{M3} in the logarithmic space.

The numerical manifestation of this structural attenuation is the “negative gradient” captured by the interaction term in Model 3. Therefore, the strong negative gradient detected in their final model is purely a statistical artifact driven by the relative decrease in the young-age sparse period associated with an increasing SP. Conversely, symptoms lacking young-age event sparsity exhibit virtually no gradient. The correlation coefficient of 0.78 underscores that this structural bias dominates the entirety of Model 3. These empirical findings definitively demonstrate that Model 3 is far from an appropriate evaluation of statistical interaction; rather, it suffers from a fatal mathematical defect that misinterprets shifts in peak onset age as vaccine-associated risk overestimations, rendering its analytical conclusions scientifically invalid.

(4) Inflation of Odds Ratios in Table 4 and the Impact of Young-Age Event Sparsity

We summarize the cumulative consequences of these methodological and mathematical breakdowns through a direct comparison between the Nagoya primary analysis and the

final model (Model 3, Table 4) of the Yaju and Tsubaki paper. A rigorous cross-model
 comparison reveals that the resulting relative odds ratios (the OR_{M3}/OR_{NS} ratio)
 exhibit an exceptionally strong negative correlation ($r = -0.83$) with EGP%. This
 empirical fact demonstrates that symptoms characterized by a lower EGP% (i.e.,
 pronounced young-age event sparsity) suffer from increasingly severe overestimation,
 with the IM_{M3}/OR_{NS} ratio ranging widely from 0.44 for Symptom 10 to 6.24 for
 Symptom 18. Structurally, Model 3 balances two opposing artifactual distortions: an
 underestimation driven by EGP deletion and an overestimation driven by the artificial
 negative gradient. Consequently, symptoms with negligible gradients (e.g., Symptoms 1
 and 10) are dominated by the EGP deletion effect, yielding IM_{M3}/OR_{NS} ratios well
 below 0.60. Conversely, young-age event-sparse symptoms (e.g., Symptoms 17, 18, 19,
 21, and 23) show marked inflation exceeding 3.0-fold due to the absolute dominance of
 the gradient-driven overestimation. Furthermore, the correlation coefficient between
 IM_{M3} and EGP% is strongly negative ($r = -0.87$). Crucially, the correlation between
 IM_{M3} and the valid baseline benchmark, OR_{NS} drops precipitously to a mere 0.27—a
 sharp, systematic decline compared to IM_{M1} (0.84) and IM_{M2} (0.38).

We now critically examine the statistical characteristics of the $SP = 0$ stratum, to which the odds ratios presented in Table 4 correspond. In this highly specific stratum, the vaccinated cohort consists solely of 29 individuals vaccinated in the year 2015—a near-vacant data cell occurring two years after the suspension of proactive recommendation—with zero corresponding unvaccinated individuals in the age 12 stratum for that specific year. Although Yaju and Tsubaki provided no methodological rationale for selecting this highly atypical stratum out of 10 available SP strata, it represents a mere 29 individuals ($<0.1\%$ of the total cohort) yet automatically guarantees the maximum possible odds ratio due to this unstated mathematical artifact. The presentation in Table 4 clearly crystallizes the analytical flaws of Model 3 as overt odds ratio overestimations.

Furthermore, while Yaju and Tsubaki devoted their discussion exclusively to “statistically significant high odds ratios” that supported their hypothesis, they offered no interpretation regarding clinically implausible estimates, such as the odds ratio of 0.34 for Symptom 10 (compared to an OR_{NS} of 0.78 in the Nagoya primary analysis), which would paradoxically imply a robust vision-protective effect of the vaccine. This highly biased, selective interpretation constitutes a clear analytical double standard and

a profound selective reporting bias, rendering the final conclusions of Table 4
scientifically invalid.

5. Discussion

The SP variable was originally introduced under the methodological guise of “excluding
pre-existing symptoms for rigorous causal inference,” serving as the foundational
architectural covariate of the Yaju and Tsubaki paper. In reality, however, it operated as
a mathematical artifact that induced severe information bias. This bias was
characterized by structurally disparate observation windows between the comparison
groups: spanning from the initial vaccination date to the study conclusion for the
vaccinated cohort, versus from age 12 to the study conclusion for the unvaccinated
cohort. Consequently, under identical SP values, unvaccinated individuals possessed
relatively longer observation periods during their youth. When coupled with the
biological reality of young-age event sparsity, this inherent information bias artificially
inflated the odds ratios for young-age event-sparse symptoms, which constituted the
underlying mathematical mechanism of their Model 2.

In Model 2, although odds ratio elevations appeared in 10 of the 24 symptoms, the magnitude of these shifts was modest, achieving statistical significance only for Symptoms 17 and 20. Because these initial findings provided insufficient evidence to substantiate a generalized risk elevation, the analytical framework was subsequently expanded to Model 3, which incorporated an SP interaction term. Under the linear relationship assumed in Model 3, the SP = 0 stratum naturally yielded the maximum possible odds ratios for young-age event-sparse symptoms. In contrast, for symptoms with high EGP% (lacking young-age sparsity), the odds ratios paradoxically decreased or flattened as SP increased. Despite this evident inconsistency in the overall trend, the authors omitted any discussion of these discrepancies, selectively presenting only the extreme peak values at SP = 0 that aligned with their hypothesis. Consequently, compared with the valid baseline benchmarks from the Nagoya primary analysis (OR_{NS}), the resulting odds ratio estimates expanded and diverged toward both extremes—upward for young-age sparse symptoms and downward for high-EGP% symptoms.

Through this sequential escalation of analytical models from Model 2 to Model 3 and ultimately to Table 4, the correlations with the valid baseline benchmark OR_{NS}

systematically diminished toward zero, while the correlations with EGP% progressively strengthened. For Symptom 18—the symptom most profoundly distorted by this artifact—the odds ratio shifted dramatically from a statistically significant lower risk of 0.70 (95% CI: 0.52–0.94) in the Nagoya primary analysis to an inflated estimate of 4.37 (95% CI: 1.63–11.61) in the Yaju and Tsubaki paper.

These methodological and mathematical findings demonstrate that the alternative models proposed by Yaju and Tsubaki are heavily compromised by structural estimation errors. To maintain the integrity of the scholarly record in applied epidemiology, secondary analytical models must be rigorously evaluated for inherent information biases and window asymmetries before drawing causal inferences.

6. Conclusions

In conclusion, this study has mathematically and epidemiologically demonstrated that the secondary analytical models and Table 4 estimates presented by Yaju and Tsubaki are artifacts driven by structural information bias and invalid mathematical extrapolation, completely lacking objective validity. Consequently, the analytical conclusions of the Yaju and Tsubaki paper must be formally re-evaluated and corrected

to reflect these methodological realities and preserve the integrity of the
epidemiological literature.

Abbreviations

EGP%: Event rate during grace period percentage

GP: Grace period

IM_M3: Increment multiplier from Model 3

OR_{M1}: Odds ratio from Model 1

OR_{M2}: Odds ratio from Model 2

OR_{M3}: Odds ratio from Model 3

OR_{NS}: Odds ratio from the Nagoya Study

SP: Study period

Declarations

Ethics approval and consent to participate

Not applicable. This study is a methodological and epidemiological evaluation based
entirely on previously published secondary data and public commentaries. No

individual participant records, administrative health datasets, or biological samples were accessed, and no fresh human participants or animals were involved; therefore, institutional ethical approval and informed consent were not required.

Consent for publication

Not applicable.

Availability of data and materials

All data analyzed during this study are derived from publicly available sources and are included within this published article, its supplementary information files, or the cited primary literature.

Competing interests

SS was involved in the primary analysis of the original Nagoya Study. However, this manuscript represents a purely retrospective, mathematical, and methodological evaluation of published open data. The author declares that he has no financial or non-financial competing interests that could inappropriately influence the collection, analysis, or interpretation of the data presented in this study.

447

448 **Funding**

449 This study received no specific grant from any funding agency in the public,
450 commercial, or not-for-profit sectors.

451

452 **Acknowledgements**

453 Not applicable.

454

455 **Authors' contributions**

456 SS conceived the study, developed the methodology, analyzed the data, generated the
457 figures, and drafted the manuscript. The author has read and approved the final
458 manuscript.

459

460 **References**

- 461 1. Suzuki S, Hosono A. No association between HPV vaccine and reported post-
462 vaccination symptoms in Japanese young women: Results of the Nagoya study.
463 Papillomavirus Res. 2018;5:96-101. doi:10.1016/j.pvr.2018.02.002.
- 464 2. 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. Jpn J Nurs Sci. 2019;16(4):433-449. doi:10.1111/jjns.12252.

3. Tsubaki H. A simultaneous adjustment analysis of two confounding factors in response to discussions on the Nagoya Cervical Cancer Immunization Study. Rinsho Hyoka (Clin Eval). 2023;50(4):485-505. https://cont.o.oo7.jp/50_4/p485-505.pdf. Accessed 10 Aug 2026. [in Japanese]

4. Suzuki S. Definition, concept, and practical example of information bias. Nihon Koshu Eisei Zasshi (Jpn J Public Health). 2026;73(6):567-573. doi:10.11236/jph.25-088. [in Japanese]

Figure Legends

Figure 1. Scatter plots illustrating the relationships between EGP% and various relative odds ratios.

a EGP% and OR_{M1}/OR_{NS} ratio. **b** EGP% and OR_{M2}/OR_{M1} ratio. **c** EGP% and IM_{M3} . **d** EGP% and OR_{M3}/OR_{NS} ratio.

Note: Closed circles (●) indicate symptoms with significantly higher rates, open circles (○) indicate symptoms with significantly lower rates, and small gray circles (●) indicate non-significant symptoms. The numbers preceded by ‘#’ indicate the nine symptoms

that showed statistically significant results.

Abbreviations:

EGP%, event rate during grace period %; IM_{M3} , increment multiplier from model 3; OR_{M1} , odds ratio from model 1; OR_{M2} odds ratio from model 2; OR_{M3} , odds ratio from model 3; OR_{NS} , odds ratio from the Nagoya Study.

Figure 2. Illustrative examples of age-matched and study period (SP)-matched comparisons for a young-age sparse symptom.

Comparison between the age-matched non-vaccinated individual (a) and the vaccinated individual (b) is indicated by black inverted triangles ($\blacktriangledown$), while comparison between the SP-matched non-vaccinated individual (c) and the vaccinated individual (b) is indicated by white inverted triangles (∇).

a Non-vaccinated individual: age = 18, SP = 6. **b** Vaccinated individual (vaccinated in 2011): age = 18, SP = 4. **c** Non-vaccinated individual: age = 16, SP = 4. (Note: Age represents the age after the birthday at the study end in 2015).

Figure 3. Variations in the impact of young-age event sparsity across different

study periods (SPs) in the unvaccinated cohort.

The diagrams schematically illustrate how the proportion of the young-age sparse period varies depending on the cumulative SP span, using $SP = 3, 6,$ and 9 as examples. The light-shaded areas represent the sparse onset period starting at age 12 (e.g., a 1-year duration), demonstrating that its relative contribution to the total observation window systematically decreases (from $1/3$ to $1/6$ and $1/9$) as the SP increases. This contraction mathematically attenuates the artificial overestimation of the odds ratios for symptoms with pronounced young-age event sparsity (e.g., Symptoms #17, #18, #19, and #23).

Figure 4. Trajectories of IM_{M3} across different study periods (SPs) for four representative symptoms.

The four linear slopes in the logarithmic space illustrate the impact of SP changes. Symptoms with high EGP% (#1 and #10) show flat trajectories with IM_{M3} values near 1.00, whereas symptoms characterized by pronounced young-age event sparsity (#17 and #18) exhibit distinct negative gradients with IM_{M3} values substantially below 1.00.

Legend: Open circle ($\circ$), Symptom #1 ($IM_{M3} = 1.00$); open square ($\square$), Symptom #10 ($IM_{M3} = 1.05$); closed circle ($\bullet$), Symptom #17 ($IM_{M3} = 0.78$); closed square ($\blacksquare$), Symptom #18 ($IM_{M3} = 0.73$).

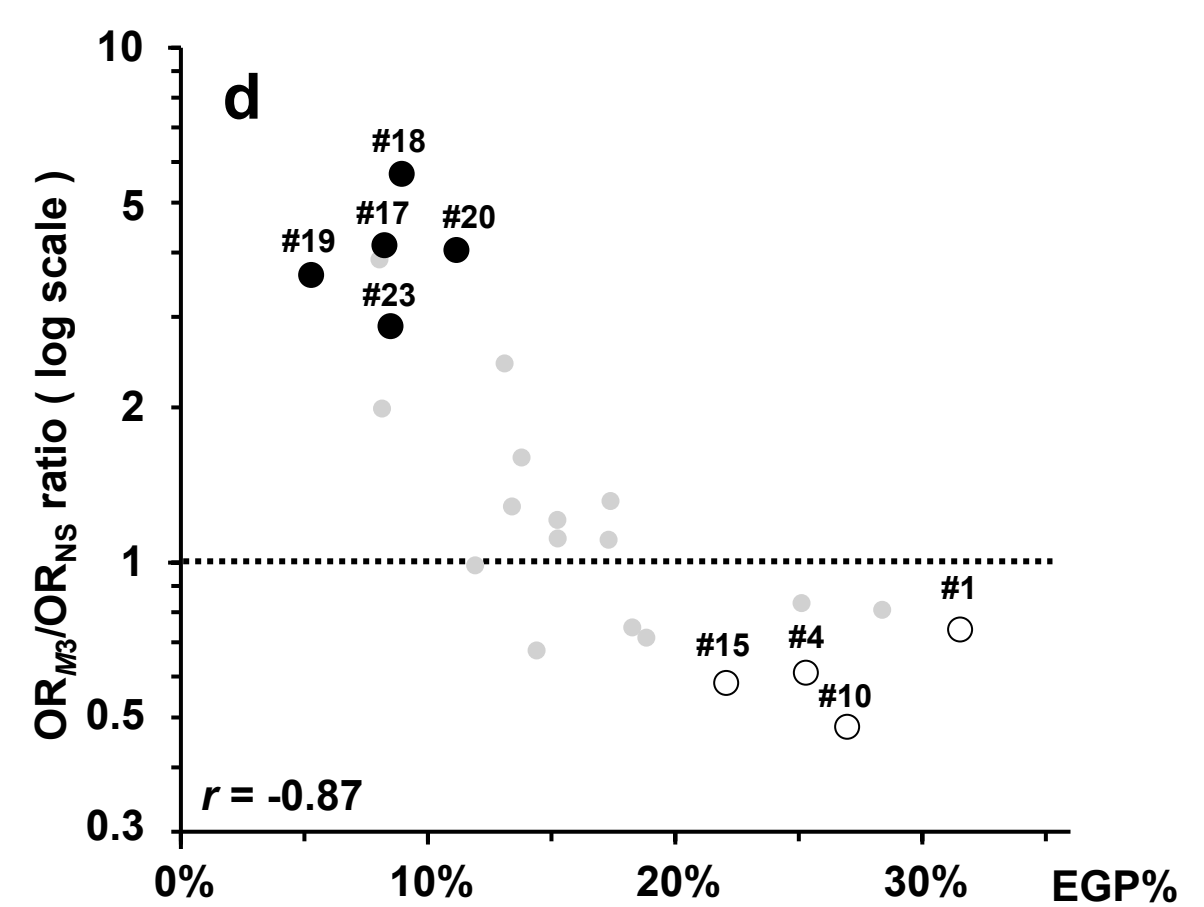
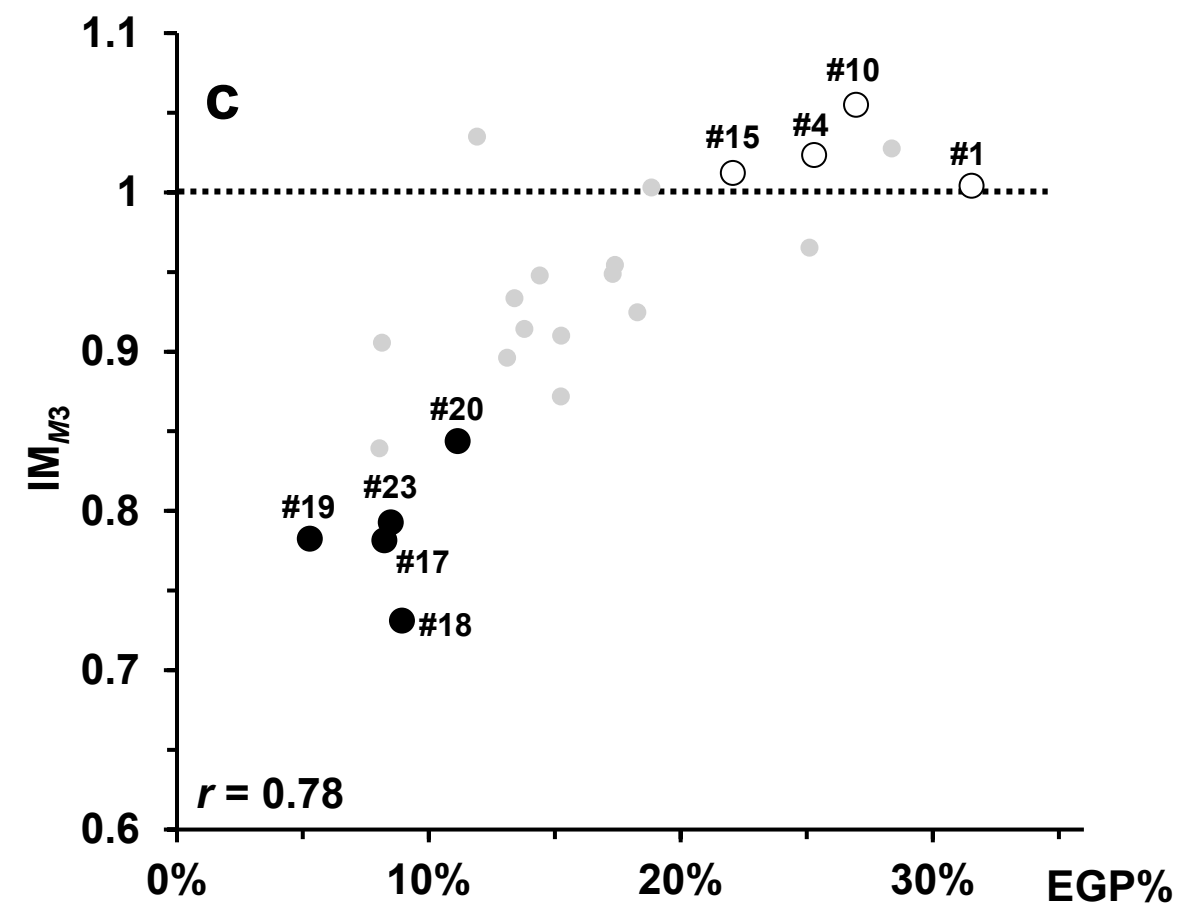
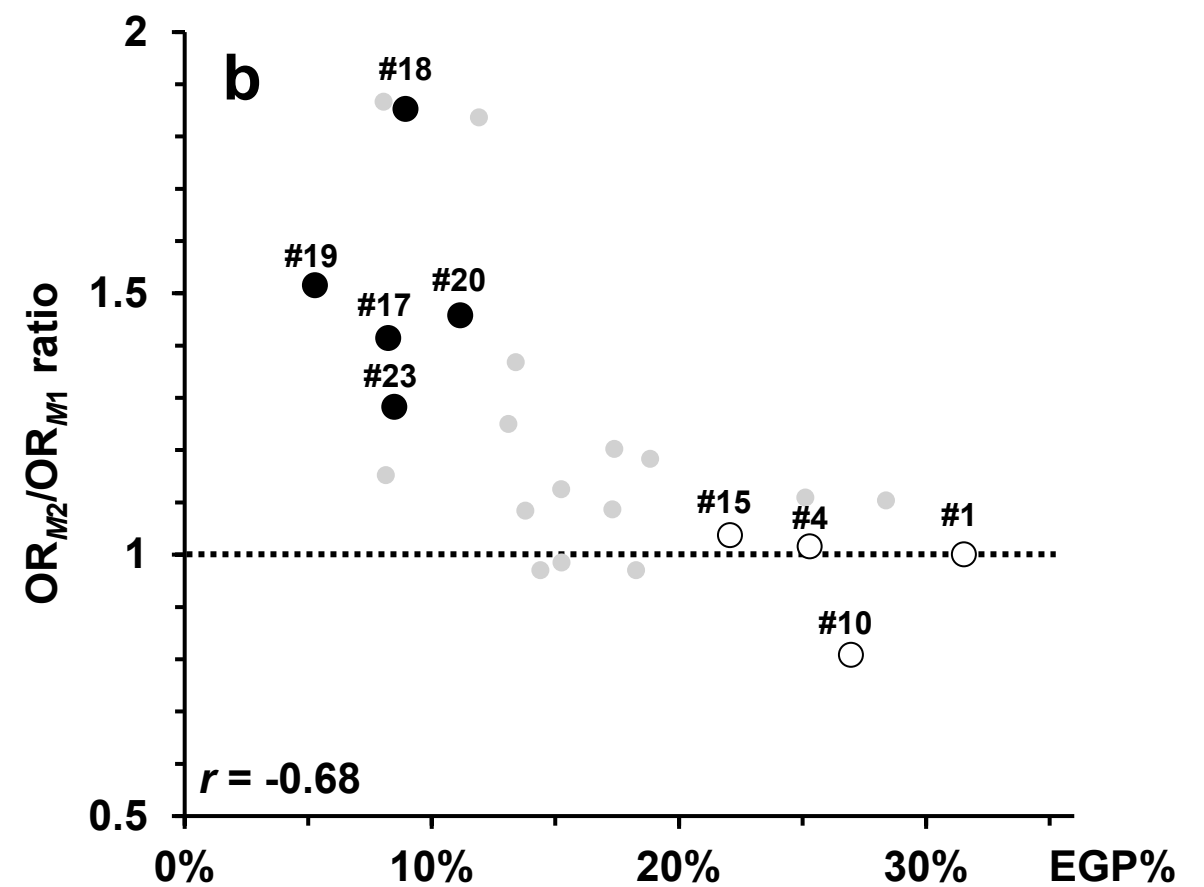
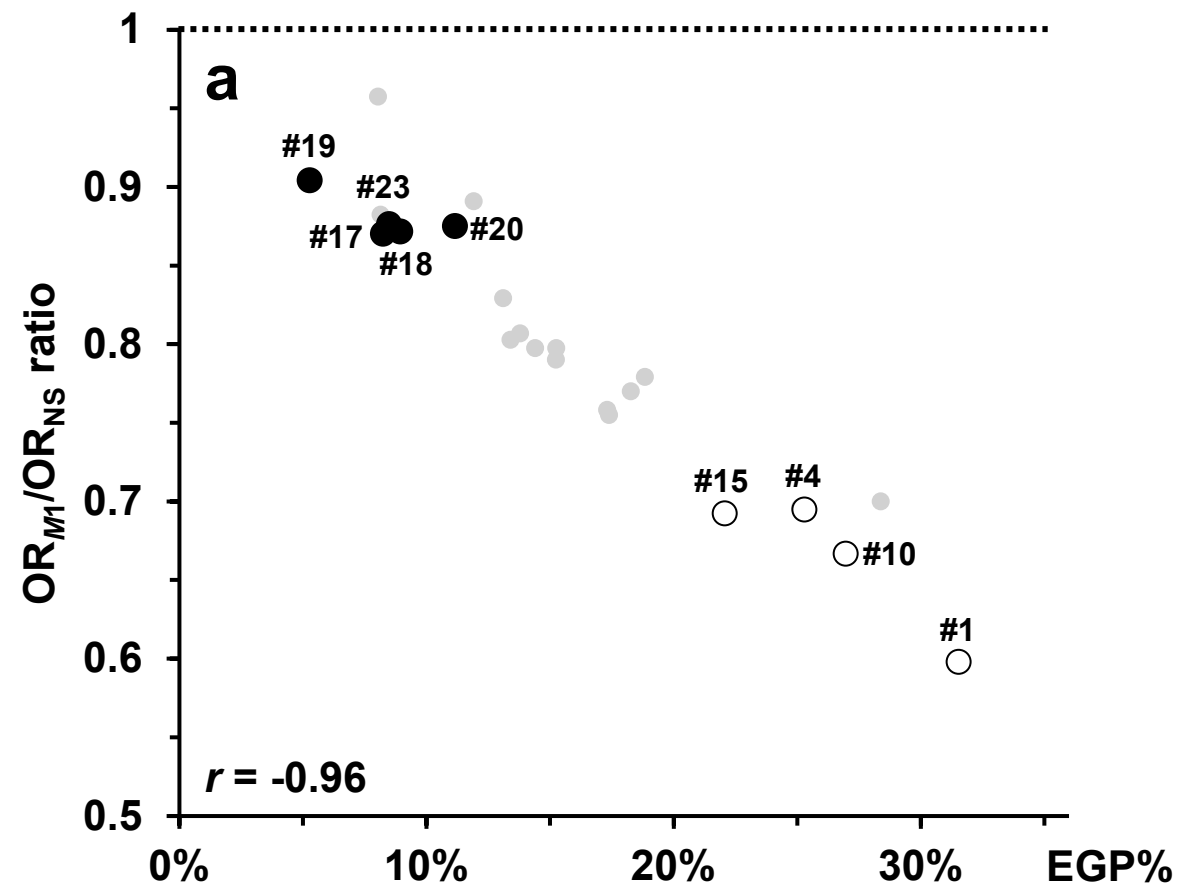


Figure 1. Scatter plots illustrating the relationships between EGP% and various relative odds ratios.

a EGP% and OR_{M1}/OR_{NS} ratio. **b** EGP% and OR_{M2}/OR_{M1} ratio. **c** EGP% and IM_{M3} . **d** EGP% and OR_{M3}/OR_{NS} ratio.

Note: Closed circles (●) indicate symptoms with significantly higher rates, open circles (○) indicate symptoms with significantly lower rates, and small gray circles (●) indicate non-significant symptoms. The numbers preceded by '#' indicate the nine symptoms that showed statistically significant results.

Abbreviations: EGP%, event rate during grace period %; IM_{M3} , increment multiplier from model 3; OR_{M1} , odds ratio from model 1; OR_{M2} , odds ratio from model 2; OR_{M3} , odds ratio from model 3; OR_{NS} , odds ratio from the Nagoya Study.

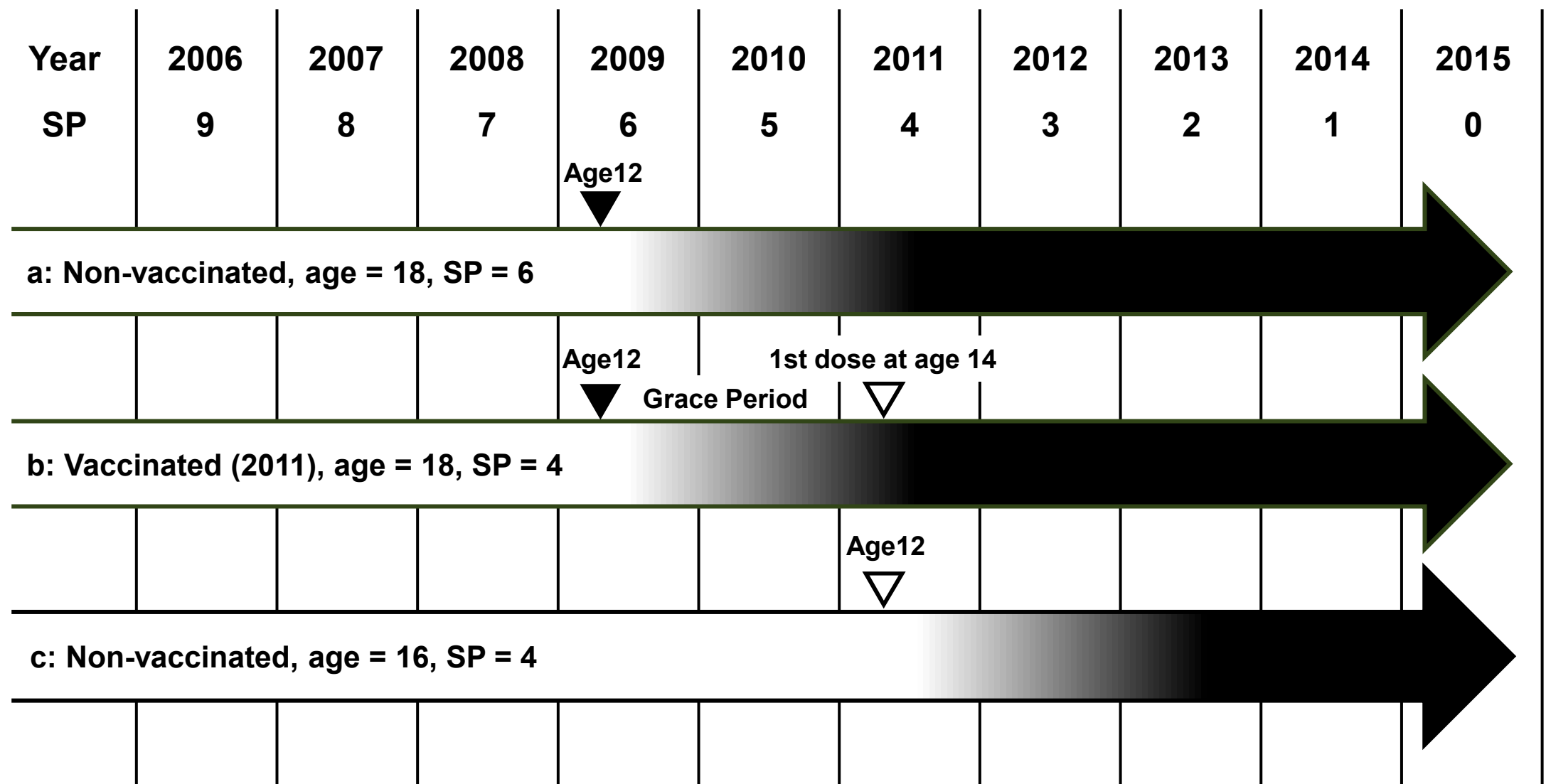


Figure 2. Illustrative examples of age-matched and study period (SP)-matched comparisons for a young-age sparse symptom.

Comparison between the age-matched non-vaccinated individual (a) and the vaccinated individual (b) is indicated by black inverted triangles (▼), while comparison between the SP-matched non-vaccinated individual (c) and the vaccinated individual (b) is indicated by white inverted triangles (▽).

a Non-vaccinated individual: age = 18, SP = 6. **b** Vaccinated individual (vaccinated in 2011): age = 18, SP = 4. **c** Non-vaccinated individual: age = 16, SP = 4. (Note: Age represents the age after the birthday at the study end in 2015).

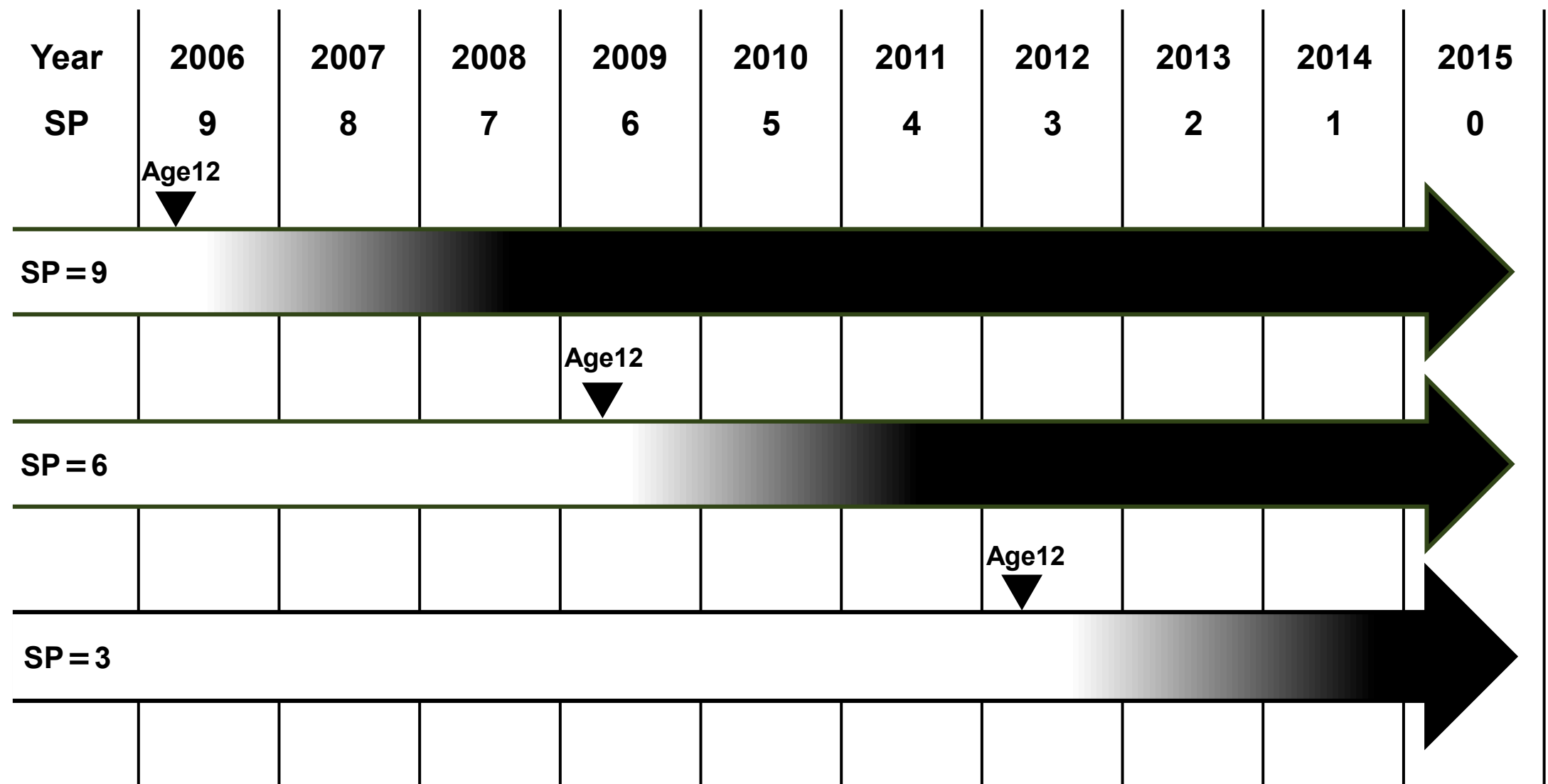


Figure 3. Variations in the impact of young-age event sparsity across different study periods (SPs) in the unvaccinated cohort.

The diagrams schematically illustrate how the proportion of the young-age sparse period varies depending on the cumulative SP span, using SP = 3, 6, and 9 as examples. The light-shaded areas represent the sparse onset period starting at age 12 (e.g., a 1-year duration), demonstrating that its relative contribution to the total observation window systematically decreases (from 1/3 to 1/6 and 1/9) as the SP increases. This contraction mathematically attenuates the artificial overestimation of the odds ratios for symptoms with pronounced young-age event sparsity (e.g., Symptoms #17, #18, #19, and #23).

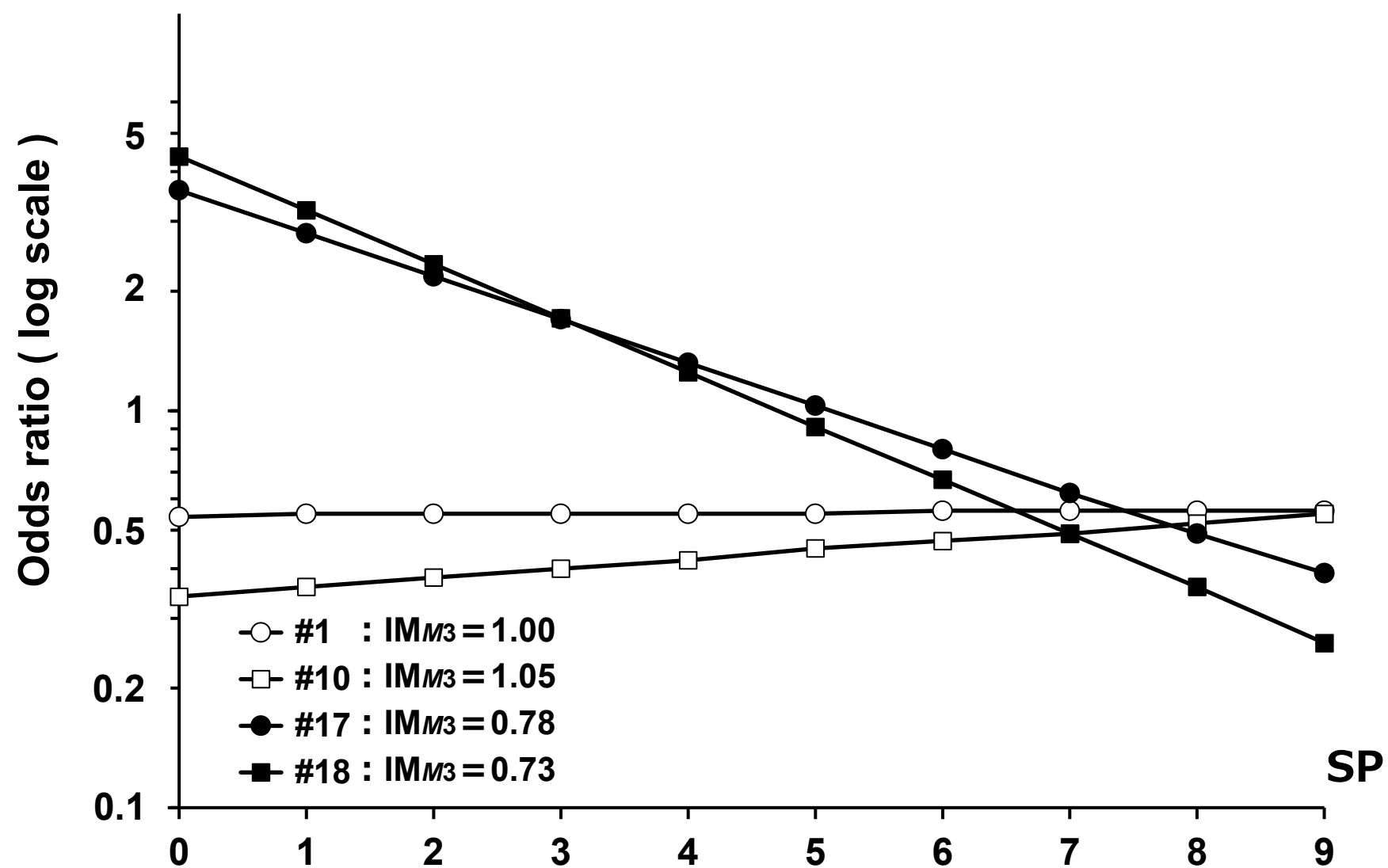


Figure 4. Trajectories of IM_{M3} across different study periods (SPs) for four representative symptoms.

The four linear slopes in the logarithmic space illustrate the impact of SP changes. Symptoms with high EGP% (#1 and #10) show flat trajectories with IM_{M3} values near 1.00, whereas symptoms characterized by pronounced young-age event sparsity (#17 and #18) exhibit distinct negative gradients with IM_{M3} values substantially below 1.00.

Legend: Open circle ($\circ$), Symptom #1 ($IM_{M3} = 1.00$); open square ($\square$), Symptom #10 ($IM_{M3} = 1.05$); closed circle ($\bullet$), Symptom #17 ($IM_{M3} = 0.78$); closed square ($\blacksquare$), Symptom #18 ($IM_{M3} = 0.73$).

Table 1: Odds ratios, EGP%, and relative odds ratios of the results of both studies, and correlation coefficients between indicators

# of symptoms		Odds ratios				EGP%	Relative odds ratios				
		Nagoya Study	Yaju-Tsubaki Paper				$\frac{OR_{M1}}{OR_{NS}}$	$\frac{OR_{M2}}{OR_{M1}}$	$\frac{OR_{M3}}{OR_{NS}}$	IM_{M3}	
			Main	Model	Model						Model
			analysis	1	2						3
Symptoms/ Correlation Coefficients		OR_{NS}	OR_{M1}	OR_{M2}	OR_{M3}		ratio	ratio	ratio		
19	Loss of ability to remember fundamental Kanji	0.73	0.66	1.00	2.82	5.3%	0.90	1.52	3.86	0.78	
21	Loss of ability to walk in a normal way	0.94	0.90	1.68	3.68	8.0%	0.96	1.87	3.91	0.84	
24	Weakness in the hands and feet	1.19	1.05	1.21	1.83	8.1%	0.88	1.15	1.54	0.91	
17	Memory decline	1.00	0.87	1.23	3.59	8.2%	0.87	1.41	3.59	0.78	
23	Sudden loss of strength	1.05	0.92	1.18	3.16	8.5%	0.88	1.28	3.01	0.79	
18	Loss of ability to do simple calculations	0.70	0.61	1.13	4.37	8.9%	0.87	1.85	6.24	0.73	
20	Involuntary uncontrollable body movement	1.20	1.05	1.53	3.19	11.2%	0.88	1.46	2.66	0.84	
22	Becoming dependent on a walking stick or wheelchair	0.55	0.49	0.90	0.77	11.9%	0.89	1.84	1.40	1.04	
8	Abnormal field of vision	0.82	0.68	0.85	1.34	13.1%	0.83	1.25	1.63	0.90	
13	Difficulty falling asleep	0.71	0.57	0.78	1.04	13.4%	0.80	1.37	1.46	0.93	
6	Poor endurance	0.88	0.71	0.77	1.12	13.8%	0.81	1.08	1.27	0.91	
7	Difficulty concentrating	0.84	0.67	0.65	0.81	14.4%	0.80	0.97	0.96	0.95	
5	Fatigue	0.81	0.64	0.72	1.27	15.2%	0.79	1.13	1.57	0.87	
11	Dizziness	0.84	0.67	0.66	0.98	15.2%	0.80	0.99	1.17	0.91	
14	Abnormally long duration of sleep	0.91	0.69	0.75	0.93	17.3%	0.76	1.09	1.02	0.95	

9	Abnormal sensitivity to light	0.98	0.74	0.89	1.08	17.4%	0.76	1.20	1.10	0.95
3	Pain in the joints or other parts of the body	0.87	0.67	0.65	0.89	18.3%	0.77	0.97	1.02	0.92
16	Hyperventilation	0.77	0.60	0.71	0.70	18.8%	0.78	1.18	0.91	1.00
15	Skin problems	0.78	0.54	0.56	0.53	22.1%	0.69	1.04	0.68	1.01
12	Cold feet	0.79	0.55	0.61	0.70	25.1%	0.70	1.11	0.89	0.97
4	Severe headache	0.95	0.66	0.67	0.61	25.3%	0.69	1.02	0.64	1.02
10	Sudden vision loss	0.78	0.52	0.42	0.34	27.0%	0.67	0.81	0.44	1.05
2	Abnormal amounts of menstrual bleeding	1.10	0.77	0.85	0.76	28.4%	0.70	1.10	0.69	1.03
1	Menstrual irregularity	0.92	0.55	0.55	0.54	31.5%	0.60	1.00	0.59	1.00
Correlation Coefficient with OR_{NS}		1.00	0.84	0.38	0.27					
Correlation Coefficient with EGP%		0.02	-0.50	-0.76	-0.83	1.00	-0.96	-0.68	-0.87	0.78

In Model 3 (Table 4) of the paper by Yaju and Tsubaki, the nine symptoms for which significant odds ratios were observed are shown in bold.

The 24 symptoms are sorted in ascending order of EGP%, which inversely reflects the degree of young-age sparsity.