medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Accurate inference methods based on the estimating equation theory for the modified Poisson and least-squares regressions

Hisashi Noma, PhD*
Department of Interdisciplinary Statistical Mathematics, The Institute of Statistical Mathematics, Tokyo, Japan
The Graduate Institute for Advanced Studies, The Graduate University for Advanced Studies (SOKENDAI), Tokyo, Japan
ORCID: http://orcid.org/0000-0002-2520-9949

Masahiko Gosho, PhD
Department of Biostatistics, Institute of Medicine, University of Tsukuba, Tsukuba, Japan

*Corresponding author: Hisashi Noma
Department of Interdisciplinary Statistical Mathematics
The Institute of Statistical Mathematics
10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
TEL: +81-50-5533-8440
e-mail: noma@ism.ac.jp

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Objectives: In clinical and epidemiological studies, the modified Poisson and least-squares regression analyses for binary outcomes have been standard multivariate analysis methods to provide risk ratio and risk difference estimates. However, their ordinary Wald-type confidence intervals can suffer from biases of the robust variance estimators and the coverage probabilities of true effect measures are substantially below the nominal level (usually 95%). To address this issue, new accurate inference methods are needed.

Methods: We provide two accurate inference methods based on the estimating equation theory for these regression models. A remarkable advantage of these regression models is that we know the correct models to be estimated: the conventional binomial regression models with log and identity links. Utilizing this modeling information, we first derive the quasi-score statistics, whose robust variances are estimated using the correct model information, and propose a confidence interval based on the regression coefficient test using χ^2-approximation. Also, to further improve the large sample approximation, we propose adapting a parametric bootstrap method to estimate the sample distribution of the quasi-score statistic using the correct model information. In addition, we developed an R package, `rqlm` (https://doi.org/10.32614/CRAN.package.rqlm), that can implement the new methods via simple commands.

Results: In extensive simulation studies, the coverage probabilities of the two new methods clearly outperformed the ordinary Wald-type confidence interval. We also illustrate the proposed methods via applications to an epidemiologic study of epilepsy. The proposed methods provided wider confidence intervals reflecting the statistical uncertainty.

Conclusions: The current standard Wald-type confidence intervals might provide misleading evidence. If erroneous evidence is reported, it can potentially influence clinical practice, public health, and policy making. These possibly inaccurate results should be circumvented through the use of effective statistical methods. The new inference methods would provide more accurate evidence in future medical studies.

Key words: generalized linear model, quasi-likelihood; estimating equation, separation problem, small sample inference.

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Logistic regression has been a standard multivariate analysis method for analyzing binary outcome data in clinical and epidemiological studies. However, the resultant odds ratio estimator cannot be directly interpreted as an effect measure; it is only be interpreted as an approximation of the risk ratio estimator when the frequency of events is small [1,2]. Thus, the use of risk ratio and risk difference is recommended as an alternative in recent guidelines; for example, the statistical reporting guidelines for *The New England Journal of Medicine* recommended avoiding using the odds ratio when reporting the results of clinical trials.

Because of the substantial limitations of logistic regression, other binomial regression models have conventionally been considered using the log or identity link functions to provide risk ratio and risk difference estimators [3]. However, the values of these binomial regression models are not limited within the range [0, 1], and the maximum likelihood (ML) estimates often cannot be defined in practice [4,5]. To address these issues, Zou [6] and Cheung [7] proposed modified Poisson and least-squares (Gaussian) regression analyses that can provide consistent risk ratio and risk difference estimators without computational difficulties. Their ideas are to formally fit the Poisson and least-squares regression models to the binary outcome data and calculate the regression coefficient estimates by the framework of the generalized linear model (GLM) [8]. The resultant estimators then provide consistent estimates of risk ratio and risk difference based on the estimating equation theory of GLM [9,10], even if the distributional assumptions are misspecified. In addition, the variance estimators should be changed to the sandwich variance estimators [11].

One relevant issue of the logistic regression is the serious biases of the regression coefficient estimator under small samples, and various correction methods have been

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

discussed [12,13]. Recently, Uno et al. [14] showed that the same bias can occur for the modified Poisson regression analysis, although this phenomenon does not occur for the modified least-squares regression. More importantly, they showed that the robust variance estimators for both of the two regression analysis methods can be biased and that the resultant Wald-type confidence intervals can seriously underestimate the actual statistical errors under small or moderate sample settings. In particular, under small and sparse data settings, these properties are especially serious, resulting in what is known as the "separation" condition—a situation where the outcome variable separates a predictor variable completely. This problem can occur even for large datasets with rare events or influential covariates. These properties can lead to seriously misleading evidence in clinical and epidemiological studies, and accurate alternative statistical inference methods are needed.

In this article, we propose new confidence intervals for the modified Poisson and least-squares regression analyses based on the estimating equation theory, especially for accurate inferences under small or moderate sample settings. A remarkable advantage of these regression models is that we know the correct models to be estimated: the conventional binomial regression models with log and identity links. Using this modeling information, we first derive the quasi-score statistics for these regression models, whose robust variances are estimated using the correct model information. The quasi-score-based inferences have been discussed for various pseudo-likelihood inferences (e.g., Mantel–Haenszel methods; [15,16]), and they are known to have favorable properties compared with the naïve Wald-type inferences. We subsequently propose a confidence interval based on the quasi-score test using χ^2-approximation. Also, to further improve small sample approximations, we propose adapting a parametric bootstrap method to estimate the sample distribution of the quasi-score statistic using the correct model

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

information. In extensive simulation studies, we show that the coverage probabilities of the two new methods clearly outperform the ordinary Wald-type confidence interval. We also illustrate the proposed methods via applications to an epidemiologic study of epilepsy. We have developed an R package, `rqlm` (https://doi.org/10.32614/CRAN.package.rqlm), that can implement the new methods via simple commands.

Modified Poisson and least-squares regressions

We consider a cohort study consisting of n participants with binary outcome $Y_1, \dots, Y_n$ ($= 1$: event occurred, $= 0$: event did not occur) and covariates $\boldsymbol{x}_i = \left(x_{i1}, x_{i2}, \dots, x_{ip}\right)^T$ for the ith subject ($i = 1, \dots, n$). Conventionally, the binomial regressions with log link and identity link functions have been considered for multivariate analyses of risk ratio and risk difference, but they involve serious theoretical difficulties for defining the ML estimates because the values of regression functions do not fall within [0, 1] [4,5]. The modified Poisson and least-squares regressions have been proposed by Zou [6] and Cheung [7] as effective methods for these multivariate analyses. Their ideas involve formally fitting the Poisson and least-squares regression models to the binary outcome data,

$$\log(\mathrm{E}[Y_i|\boldsymbol{x}_i]) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}$$

$$\mathrm{E}[Y_i|\boldsymbol{x}_i] = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}$$

The resultant quasi-ML estimators $\widehat{\boldsymbol{\beta}}$ of the regression coefficients $\boldsymbol{\beta} = \left(\beta_0, \beta_1, \dots, \beta_p\right)^T$ become consistent estimators of the log-transformed risk ratios and risk differences on the target population [14]. The principle of these estimating methods is founded on the estimating equation theory of GLM [10]; that is, the estimating functions are unbiased even if the distribution forms are misspecified as long as the functional forms of the regression functions are correctly specified. In particular, for the modified least-

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

squares regression, the Gauss–Markov theorem [17,18] is adapted and the quasi-ML estimator becomes the best linear unbiased estimator (BLUE); $\widehat{\boldsymbol{\beta}}$ is an unbiased estimator and the most precise estimator for the regression coefficients of the binomial regression function [14]. The standard errors of $\widehat{\boldsymbol{\beta}}$ of both models are consistently estimated by the sandwich variance estimator [11].

Confidence intervals for risk ratio and risk difference

Confidence intervals based on the quasi-score statistics

The modified Poisson and least-squares regressions are effective methods for multivariate analyses of risk ratio and risk difference; however, their ordinary Wald-type confidence intervals can seriously underestimate the statistical errors under small or moderate sample settings [14]. To address these issues, we first derive quasi-score tests for the regression coefficients. The two models are formulated as specific cases of the GLM, and the quasi-likelihood estimating functions are expressed as

$$U(\boldsymbol{\beta}) = \sum_{i=1}^{n} \boldsymbol{D}_i^T V_i^{-1}(Y_i - \mu_i) = 0$$

where μ_i is the mean function ($= \exp(\boldsymbol{\beta}^T \boldsymbol{x}_i)$ for the Poisson model and $= \boldsymbol{\beta}^T \boldsymbol{x}_i$ for the Gaussian model) and $\boldsymbol{D}_i = \partial \mu_i / \partial \boldsymbol{\beta}$; also, $V_i = v(\mu_i)$ is the variance function of the outcome variable ($= \mu_i$ for the Poisson model and $= 1$ for the Gaussian model) ($i = 1, \ldots, n$). If the variance functions are correctly specified, the covariance matrices of $U(\boldsymbol{\beta})$ become the Fisher information matrices $I(\boldsymbol{\beta}) = -E[\partial U(\boldsymbol{\beta}) / \partial \boldsymbol{\beta}]$. However, the variance functions are misspecified for these cases; thus, they become the robust covariance matrices, $J(\boldsymbol{\beta}) = E[U(\boldsymbol{\beta}) U^T(\boldsymbol{\beta})]$. Also, the special feature of these inferences is that the correct models are known for us and the expectations can be substituted into the binomial regression models. Then, the concrete forms of the models

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

are then expressed for

$$U(\boldsymbol{\beta}) = \sum_{i=1}^{n} \boldsymbol{x}_i(Y_i - \mu_i) = \boldsymbol{X}(\boldsymbol{Y} - \boldsymbol{\mu})$$

$$J(\boldsymbol{\beta}) = \sum_{i=1}^{n} V[Y_i]\boldsymbol{x}_i\boldsymbol{x}_i^T = \sum_{i=1}^{n} \mu_i(1-\mu_i)\boldsymbol{x}_i\boldsymbol{x}_i^T = \boldsymbol{X}\boldsymbol{W}\boldsymbol{X}^T$$

where $\boldsymbol{Y} = (Y_1, \dots, Y_n)^T, \boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T, \boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$, and $\boldsymbol{W} = \mathrm{diag}\{\mu_1(1-\mu_1), \dots, \mu_n(1-\mu_n)\}$. Although the definitions of the mean function μ_i differ between the two models, the function forms are the same. Note that $\mu_1, \dots, \mu_n$ should be truncated on [0, 1] on $\boldsymbol{W}$ because the individual variance functions substantially estimate the variances of binomial variables and should not be negative values. However, those in $U(\boldsymbol{\beta})$ should not be truncated; if they are truncated, the quasi-score functions are biased and unrealistically singular results can be obtained. We also note that $E[U(\boldsymbol{\beta})] = \boldsymbol{0}$ and $V[U(\boldsymbol{\beta})] = J(\boldsymbol{\beta})$ without large sample approximations. The quasi-score test statistics for the joint null hypotheses H0: $\boldsymbol{\beta} = \boldsymbol{\beta}_{\mathrm{null}}$ are constructed using the exact means and covariance matrices. However, these null hypotheses are usually outside the scope of interest in practice.

We consider tests for composite null hypotheses H0: $\beta_1 = \beta_{1,\mathrm{null}}$ that correspond to the hypothesis tests for individual risk ratio and risk difference; without loss of generality, we consider the tests of regression coefficients of first explanatory variables and denote $\boldsymbol{\beta}_c = (\beta_2, \beta_3, \dots, \beta_p)^T$. The quasi-score statistics are then constructed as

$$T(\beta_{1,\mathrm{null}}) = U^T(\tilde{\boldsymbol{\beta}})J^{-1}(\tilde{\boldsymbol{\beta}})U(\tilde{\boldsymbol{\beta}})$$

where $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_0, \beta_{1,\mathrm{null}}, \tilde{\boldsymbol{\beta}}_c)^T$; $\tilde{\beta}_0$ and $\tilde{\boldsymbol{\beta}}_c = (\tilde{\beta}_2, \tilde{\beta}_3, \dots, \tilde{\beta}_p)^T$ are the constrained quasi-ML estimates of $\{\beta_0, \boldsymbol{\beta}_c\}$ under H0. The constrained quasi-ML estimates can be calculated via the same modified Poisson and least-squares regression analyses when x_{1i} is dropped from the explanatory variables and offsets $\beta_{1,\mathrm{null}}x_{1i}$ are added ($i = 1, \dots, n$).

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Under the null hypotheses, the quasi-score test statistics follow the χ^2-distribution with approximately one degree of freedom [19]. Also, using these quasi-score tests, we can construct the $100 \times (1-\alpha)\%$ confidence intervals of β_1 by the sets of null values that satisfy

$$T(\beta_{1,\text{null}}) \leq \chi^2_{1,1-\alpha}$$

where $\chi^2_{1,1-\alpha}$ is the upper αth percentile of the χ^2-distribution with one degree of freedom. The confidence limits can be calculated by adequate numerical methods (e.g., the bisectional methods [20]).

As shown in the simulation studies, the confidence intervals based on the quasi-score statistics have generally favorable properties compared with the ordinary Wald-type confidence intervals obtained using the standard sandwich variance estimators. However, the confidence intervals when using the large sample χ^2 approximations still have limitations with respect to achieving sufficient coverage performance under small sample settings [21]; in addition, more accurate approximations of the sample distributions are needed for valid inferences.

Bootstrap confidence intervals based on the quasi-score statistics

To improve the accuracy of approximations of the sample distributions of the quasi-score statistics, we propose using the bootstrap method. Again, we focus on the advantage of these regression methods—that the correct distributional assumptions on the target population are known (binomial regression models). Thus, we propose performing parametric bootstrap resampling from the "correct" binomial regression models by substituting the regression coefficients $\boldsymbol{\beta}$ for the null value and the constrained quasi-ML estimates. The bootstrap algorithm for the tests of H0: $\beta_1 = \beta_{1,\text{null}}$ is as follows.

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Algorithm (bootstrap tests for the quasi-score statistics)

1. For the modified Poisson and least-squares regression models, compute the constrained quasi-ML estimates $\tilde{\beta}_0$ and $\tilde{\boldsymbol{\beta}}_c$ under H0: $\beta_1 = \beta_{1,\text{null}}$.

2. Resample $Y_1^{(b)}, \dots, Y_n^{(b)}$ from the binomial regression models with log or identity links whose regression coefficients $\boldsymbol{\beta}$ are fixed to $\tilde{\boldsymbol{\beta}}$,

$$\Pr\,(Y_i = 1|\boldsymbol{x}_i) = \exp\,(\tilde{\beta}_0 + \beta_{1,\text{null}}x_{i1} + \tilde{\beta}_2 x_{i2} + \cdots + \tilde{\beta}_p x_{ip})$$

$$\Pr\,(Y_i = 1|\boldsymbol{x}_i) = \tilde{\beta}_0 + \beta_{1,\text{null}}x_{i1} + \tilde{\beta}_2 x_{i2} + \cdots + \tilde{\beta}_p x_{ip}$$

via parametric bootstraps, B times ($i = 1, \dots, n$; $b = 1,2, \dots, B$). Note that, if the values of regression functions on the right-hand sides of these equations exceed the range [0, 1], they should be truncated at 0 or 1. Also, the design matrix $\boldsymbol{X}$ is not altered from the original data across the resampling.

3. Compute the quasi-score statistic $T^{(b)}(\beta_{1,\text{null}})$ for the b th bootstrap sample $Y_1^{(b)}, \dots, Y_n^{(b)}$ ($b = 1,2, \dots, B$).

4. Calculate the empirical distribution function of $T^{(1)}(\beta_{1,\text{null}}), \dots, T^{(B)}(\beta_{1,\text{null}})$ — specifically, $\bar{F}_{T(\beta_{1,\text{null}})}(t)$, which is the bootstrap estimate of the sample distribution of $T(\beta_{1,\text{null}})$.

5. Implement the hypothesis test for $T(\beta_{1,\text{null}})$ using $\bar{F}_{T(\beta_{1,\text{null}})}(t)$ for the reference distribution.

Because accurate tail area estimation of the null distribution requires a large number of replications [22], the number of bootstrap resamplings B should be sufficiently large (usually, at least 1000).

The corresponding $100 \times (1-\alpha)\%$ confidence intervals of β_1 can be constructed by the sets of $\beta_{1,\text{null}}$ that satisfy

$$T(\beta_{1,\text{null}}) \leq \bar{F}_{T(\beta_{1,\text{null}})}(1-\alpha)$$

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

The confidence limits can also be calculated by adequate numerical methods (e.g., the bisectional methods; [20]). The parametric bootstrap approach effectively uses the distributional information of the correct models; thus, the approximation is expected to be improved compared with the naïve asymptotic normal approximation. The actual performances are demonstrated in the simulation studies.

Software

We developed an R package, `rqlm` (https://cran.r-project.org/web/packages/rqlm), to perform all of the proposed methods via simple commands.

Simulations

To illustrate the operating characteristics of the proposed methods, we performed simulation studies. For data generation, we considered the binomial regression models with log and identity link functions; parameter settings were selected to mimic the epilepsy epidemiologic study described in the next section. Four explanatory variables were considered: x_{i1} was the main treatment/exposure variable; it followed a Bernoulli distribution with probability 0.20 or 0.10; x_{i2} was a confounding variable that followed a Bernoulli distribution with probability 0.773 and was found to have a correlation with x_{i1} through measurement of the odds ratio [OR] $\Pr(x_{i1}=1)\Pr(x_{i2}=0)/\Pr(x_{i1}=0)\Pr(x_{i2}=1)$ = 25, 15, and 5; x_{i3} followed a Bernoulli distribution with probability 0.455, and x_{i4} followed $N(29.0,\ 7.37)$. The outcome variable Y_i was then generated from a Bernoulli distribution with probability

$$\Pr(Y_i=1)=\exp(\beta_0+\beta_1x_{i1}+\beta_2x_{i2}+\beta_3x_{i3}+\beta_4x_{i4})$$

$$\Pr(Y_i=1)=\beta_0+\beta_1x_{i1}+\beta_2x_{i2}+\beta_3x_{i3}+\beta_4x_{i4}$$

The intercept β_0 was set by controlling the overall event rate of the cohort; the event rate

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

was varied as 0.40 and 0.20. The other regression coefficients were set as $(\beta_1, \beta_2, \beta_3, \beta_4) = (0.205, -0.271, 0.000, 0.153)$ for the former model, and $= (0.116, -0.041, 0.0037, 0.023)$ for the latter model. We considered the sample size $n = 20, \ldots, 100$. We performed 5,000 simulations for the 108 scenarios, for both the risk ratio and risk difference regression models.

For comparisons, we analyzed the individual dataset by the ordinary modified Poisson and least-squares regressions using the Wald-type confidence intervals obtained by the standard sandwich variance estimator. We then applied the two proposed confidence intervals based on the quasi-score statistic and the bootstrap approach. For the latter method, we performed 2,000 bootstrap resamplings to estimate the bootstrap distributions. We assessed the coverage probabilities of 95% confidence intervals for β_1 of the three methods.

The results of the simulation studies are presented in Figures 1 and 2 for the modified Poisson regression and in Figures 3 and 4 for the modified least-squares regression. The empirical coverage rates of 95% confidence intervals for the 5,000 simulations are plotted. For the modified Poisson regression, the Wald-type 95% confidence intervals generally suffered undercoverage under small or moderate sample settings. These results were caused by the biases of both of the regression coefficient estimates and the robust standard error estimates. The degrees of undercoverage became serious if the event rate or the exposure rate became too small and the correlations of x_{i1} and x_{i2} became too large; these settings possibly coincide with the separation or quasi-separation settings. In addition, the two proposed confidence intervals retained the coverage probabilities around the nominal level (95%) for most of the scenarios. The quasi-score-based confidence interval could reflect undercoverage bias under some small sample settings (generally, $n \leq 50$); however, the coverage properties were consistently favorable

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

compared with those of the Wald-type confidence interval. In addition, the bootstrap-based confidence interval showed better coverage probabilities in general, as expected. Even for especially small sample settings ($n = 20$), the coverage probabilities had values around the nominal level (95%) consistently.

For the modified least-squares regression, the overall features of the results of simulations were similar to those for the modified Poisson regression cases. The Wald-type 95% confidence intervals were generally prone to undercoverage bias under small or moderate samples. Note that the quasi-ML estimates correspond to the BLUE according to the Gauss–Markov theorem; thus, there were no biases for the point estimator. However, the robust variance estimator was seriously biased under small or moderate sample settings. The degrees of undercoverage also became serious if the event rate or the exposure rate became excessively small and the correlations of x_{i1} and x_{i2} became large. In addition, the proposed confidence intervals showed favorable coverage properties. The quasi-score-based confidence interval could be prone to undercoverage under some settings; however, the coverage properties were much better than those of the Wald-type confidence interval. Also, the bootstrap-based confidence interval further improved the coverage properties consistently. Even for especially small sample settings ($n = 20$), the coverage probabilities consistently have values close to the nominal level (95%). These simulation results would indicate the proposed bootstrap-based confidence interval provide accurate interval estimates generally, and the validity is retained even under small and sparse data settings.

Applications

To illustrate the usefulness of the proposed methods, we applied them to an epidemiologic study of epilepsy by Arai et al. [23], which is a retrospective cohort study to evaluate the

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

factors associated with the employment statuses of patients with a history of childhood-onset drug-resistant epilepsy ($N = 56$). We analyzed this cohort data using the modified Poisson and least-squares regressions. The outcome was employment status (1 = nonemployment, 0 = employment; the number of events was 14), and four explanatory variables were involved: age at follow up, gender, mood disorder symptoms, and graduating from a school for special needs education. Note that the last two variables were highly correlated, and both of these variables were found to be strongly associated with the outcome through univariate analyses [23].

The results are presented in Table 1. The confidence intervals based on the quasi-score statistic and bootstrap approach were asymmetric around the quasi-ML estimates for many cases. In some cases, the locations of the confidence intervals also differed substantially from those of the Wald-type confidence intervals. This phenomenon is generally known to occur in the case of efficient score-based confidence intervals [21]. For the modified Poisson regression, only the 95% Wald-type confidence interval of the fourth covariate did not cover the null value (= 1). The proposed two confidence intervals also did not cover 1; however, the bootstrap confidence interval showed a large effect compared with the Wald-type confidence interval. For the third covariate, the confidence interval based on the quasi-score statistic was narrower than the Wald-type confidence interval and did not cover 1; however, the bootstrap-based confidence interval was much wider and did not involve 1. These results might be influenced by the strong correlation between these two covariates. In addition, for the modified least-squares regression, the Wald-type confidence interval of the third covariate was substantially widened by the bootstrap-based approach. Also, the bootstrap confidence interval for the fourth covariate was wider than the Wald-type confidence interval. Considering the operating characteristics shown by simulations, the conclusions obtained from the ordinary Wald-

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

type confidence intervals might be misleading; the improved methods would likely provide more precise evidence.

Discussion

The modified Poisson and least-squares regressions have been widely used in recent clinical and epidemiological studies because they can provide interpretable effect measure estimates without computational difficulties. Considering the difficulty of the interpretability of the odds ratio, these methods will be increasingly adopted in future studies as effective alternatives to the conventional logistic regression. In this article, we revealed that the ordinary Wald-type confidence intervals of these methods are prone to serious undercoverage under small or moderate sample settings, especially under separation or quasi-separation settings. As clearly shown in the real data example, the ordinary confidence intervals might provide misleading evidence; if erroneous evidence is reported, it can potentially influence clinical practice, public health, and policy making. These possibly inaccurate results should be circumvented through the use of effective statistical methods. The new accurate confidence intervals and their numerical evidence provided in this article will be useful in future medical studies.

In the simulation-based evidence, the ordinary Wald-type confidence interval showed serious undercoverage performance. For the modified Poisson regression, a relevant reason for the invalid property is the small sample bias; although the bias of the ML estimator for the GLM is well known [24], a similar bias can occur for the quasi-ML estimator for the misspecified models. Effective solutions to address this bias would be the Firth-type bias correction [25] and higher-order bias correction methods [24]. We will discuss these bias reduction methods in subsequent papers. In addition, for the modified least-squares regression, the resultant estimator is unbiased because it corresponds to the

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

BLUE according to the Gauss–Markov theorem and no corrections are needed if this criterion is considered. However, the undercoverage properties shown in the simulation studies were severe and the bias of the ordinary robust variance estimator should be adequately addressed in practice. The proposed new methods are expected to be a good solution.

Another future issue is adapting improved robust variance estimators for the Wald-type confidence interval [26]. Although the accuracies of existing improved robust variance estimators are generally better than that of the ordinary sandwich variance estimator, deterministic conclusions cannot be provided for their relative performances generally [26] because all the methods are founded on approximations (e.g., higher-order approximations). Although simulation-based numerical evidence can provide good case-by-case comparative performance, they would not be generic properties. However, the bootstrap-based approach proposed in the present article is founded on the quasi-score statistic that effectively uses information of the null hypothesis and adapts the flexible sample distribution estimate. We therefore believe it is one of the most effective methods among the competing methods. The proposed methods can be used as accurate and effective alternatives to the ordinary Wald-type inference methods in clinical and epidemiological studies.

Acknowledgements

The authors would like to thank Dr. Y. Arai (Tottori University) for permission to use the valuable data and K. Nakazono (The Institute of Statistical Mathematics) for his helpful comments on the earlier draft.

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Competing interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Research Funding

This study was supported by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (grant numbers: JP23K11931, JP22H03554, JP24K21306, and JP23H03063).

Data availability

R package for implementing the proposed methods is available at CRAN (https://cran.r-project.org/web/packages/rqlm).

References

1. Greenland S. Interpretation and choice of effect measures in epidemiologic analysis. Am J Epidemiolo 1987;125:761-768.
2. Nurminen M. To use or not to use the odds ratio in epidemiologic analyses. Eur J Epidemiol 1995;11:365-371.
3. Rothman KJ, Greenland G, Lash TL. Modern epidemiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008.
4. McNutt LA, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol 2003;157:940-3.
5. Wallenstein S, Bodian C. Epidemiologic programs for computers and calculators. Inferences on odds ratios, relative risks, and risk differences based on standard regression programs. Am J Epidemiol 1987;126:346-55.

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

6. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159:702-6.
7. Cheung YB. A modified least-squares regression approach to the estimation of risk difference. Am J Epidemiol 2007;166:1337-44.
8. Nelder JA, Wedderburn RWM. Generalized linear models. J Royal Stat Soc A 1972;135:370-384.
9. Godambe VP, Heyde CC. Quasi-likelihood and optimal estimation. Int Stat Rev 1987;55:231-244.
10. Wedderburn RWM. Quasi-likelihood functions, generalized linear models, and the gauss-newton method. Biometrika 1974;61:439-447.
11. White H. Maximum likelihood estimation of misspecified models. Econometrica 1982;50:1-25.
12. Albert A, Anderson JA. On the existence of the maximum likelihood estimates in logistic regression models. Biometrika 1984;71:1-10.
13. Zorn C. A solution to separation in binary response models. Polit Anal 2005;13:157-170.
14. Uno S, Noma H, Gosho M. Firth-type penalized methods of the modified poisson and least-squares regression analyses for binary outcomes. Biom J 2024;66:e202400004.
15. Sato T. Confidence limits for the common odds ratio based on the asymptotic distribution of the mantel-haenszel estimator. Biometrics 1990;46:71-80.
16. Noma H, Nagashima K. A note on the mantel-haenszel estimators when the common effect assumptions are violated. Epidemiol Meth 2016;5:19-35.
17. Plackett RL. A historical note on the method of least squares. Biometrika 1949;36:458-460.
18. David FN, Neyman J. Extension of the markoff theorem on least squares. Stat Res

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Mem 1938;2:105-116.

19. Cox DR, Hinkley DV. Theoretical statistics. London: Chapman and Hall; 1974.
20. Burden RL, Faires JD. Numerical analysis. 9th ed ed. Boston: Cengage Learning; 2011.
21. Noma H. Confidence intervals for a random-effects meta-analysis based on bartlett-type corrections. Stat Med 2011;30:3304-3312.
22. Efron B, Tibshirani R. An introduction to the bootstrap. New York: CRC Press; 1994.
23. Arai Y, Okanishi T, Noma H, et al. Prognostic factors for employment outcomes in patients with a history of childhood-onset drug-resistant epilepsy. Front Pediatr 2023;11:1173126.
24. Cordeiro GM, McCullagh P. Bias correction in generalized linear models. J Royal Stat Soc B 1991;53:629-643.
25. Firth D. Bias reduction of maximum likelihood estimates. Biometrika 1993;80:27-38.
26. Gosho M, Ishii R, Noma H, Maruo K. A comparison of bias-adjusted generalized estimating equations for sparse binary data in small-sample longitudinal studies. Stat Med 2023;42:2711-2727.

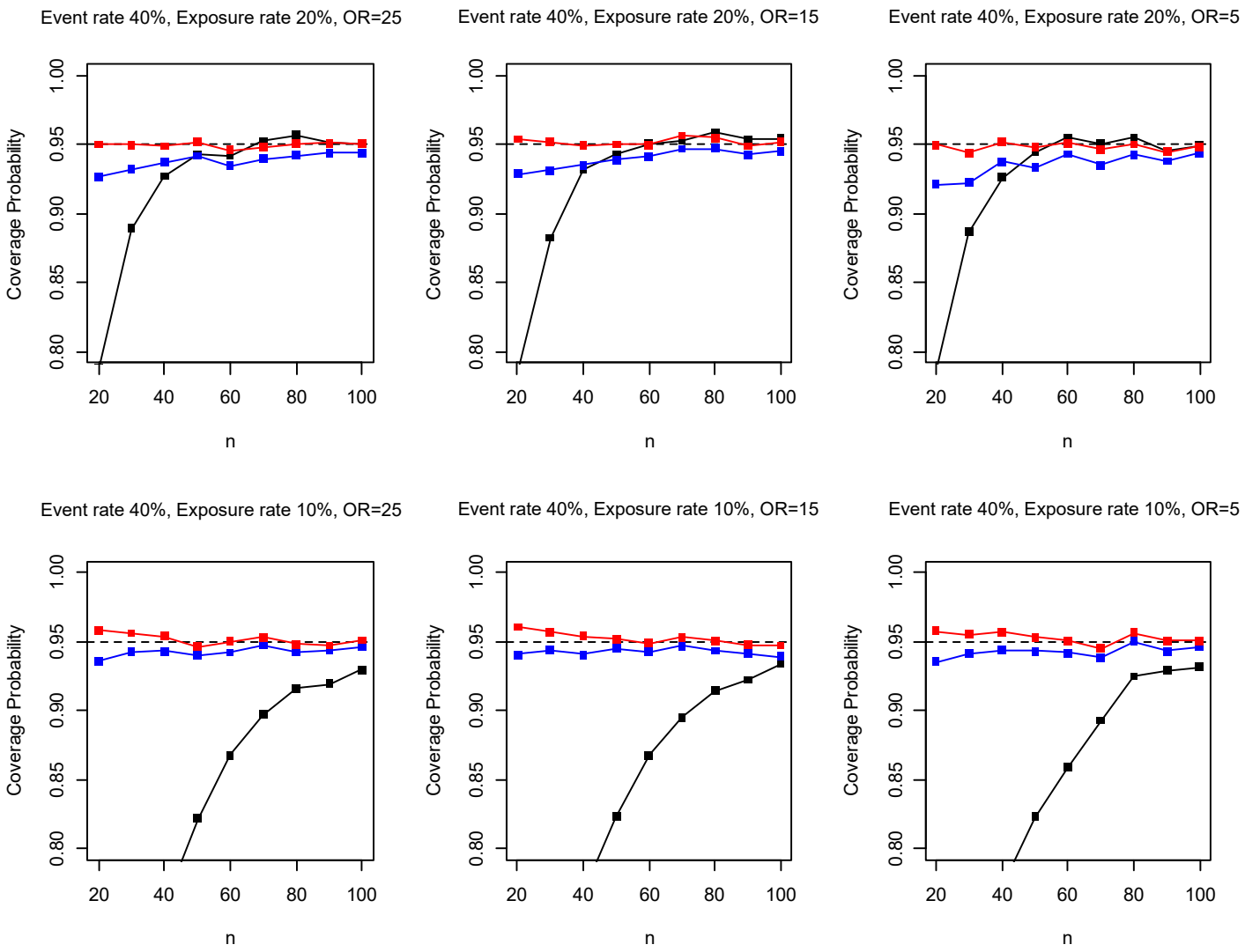

Figure 1. Results of the simulations for the modified Poisson regression (**black**: ordinary Wald CI, **blue**: quasi-score CI, **red**: bootstrap CI by the quasi-score statistic; CI: confidence interval).

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (**which was not certified by peer review**) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

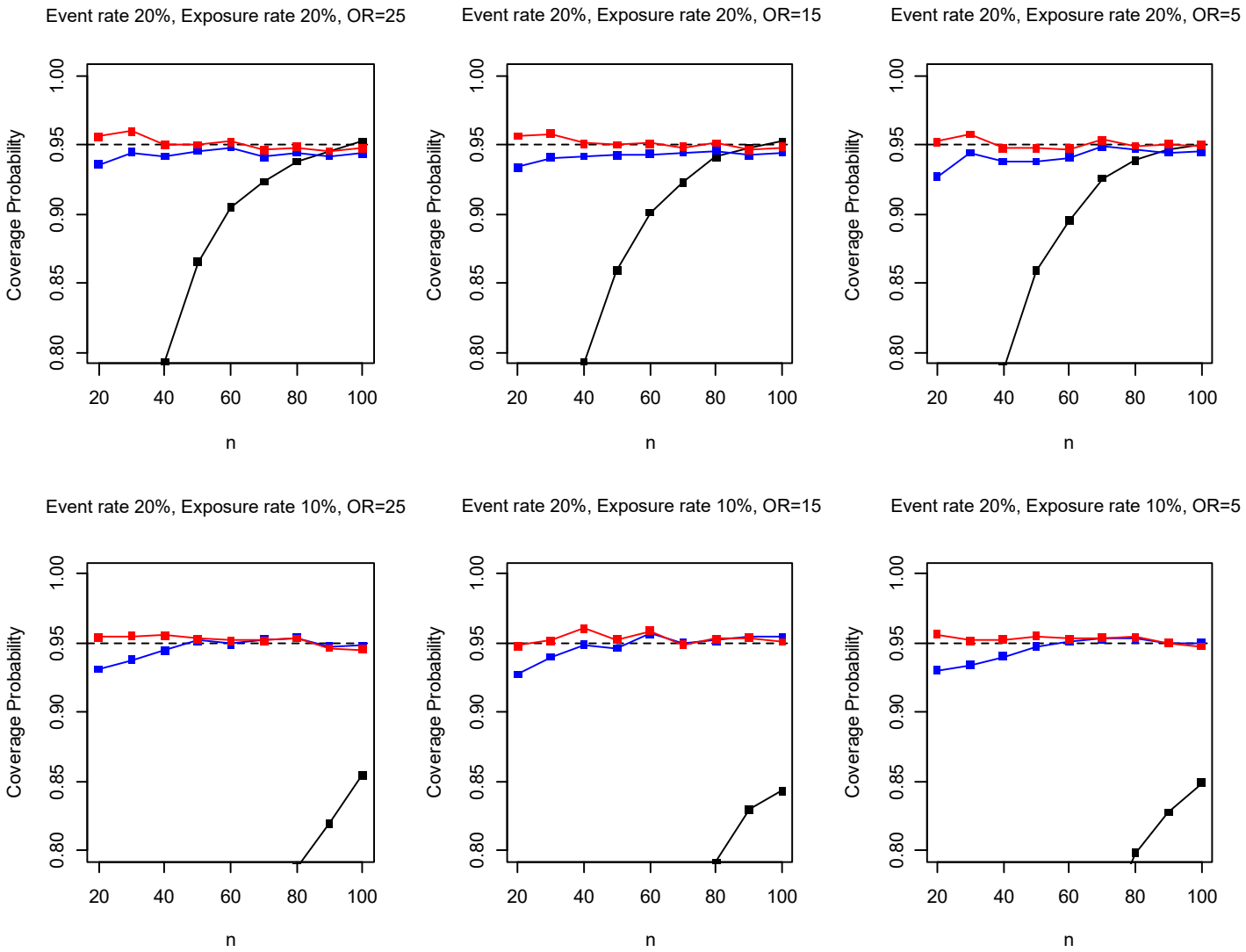

Figure 2. Results of the simulations: modified Poisson regression (**black**: ordinary Wald CI, **blue**: quasi-score CI, **red**: bootstrap CI by the quasi-score statistic; CI: confidence interval).

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (**which was not certified by peer review**) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

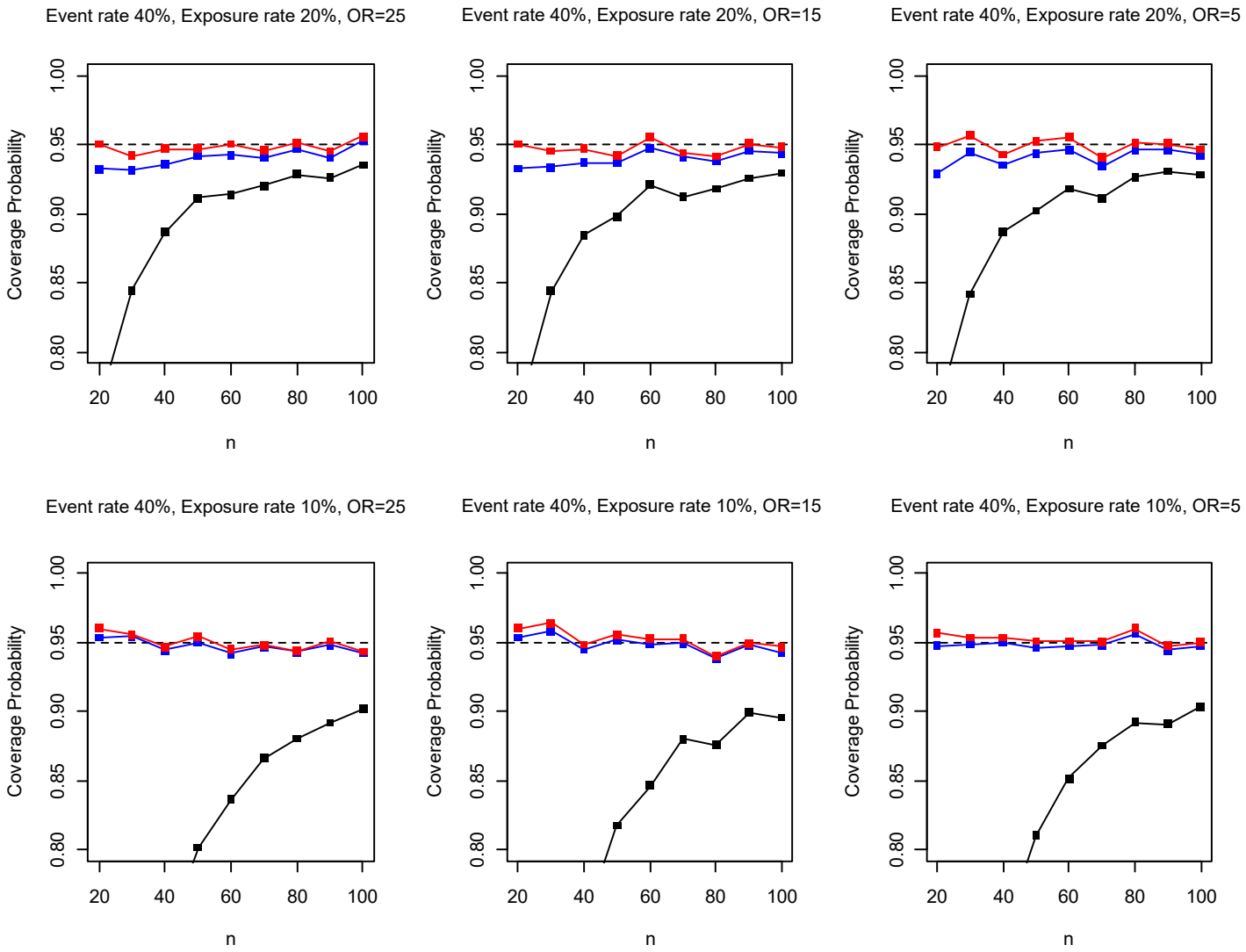

Figure 3. Results of the simulations: modified least-squares regression (**black**: ordinary Wald CI, **blue**: quasi-score CI, **red**: bootstrap CI by the quasi-score statistic; CI: confidence interval).

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (**which was not certified by peer review**) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

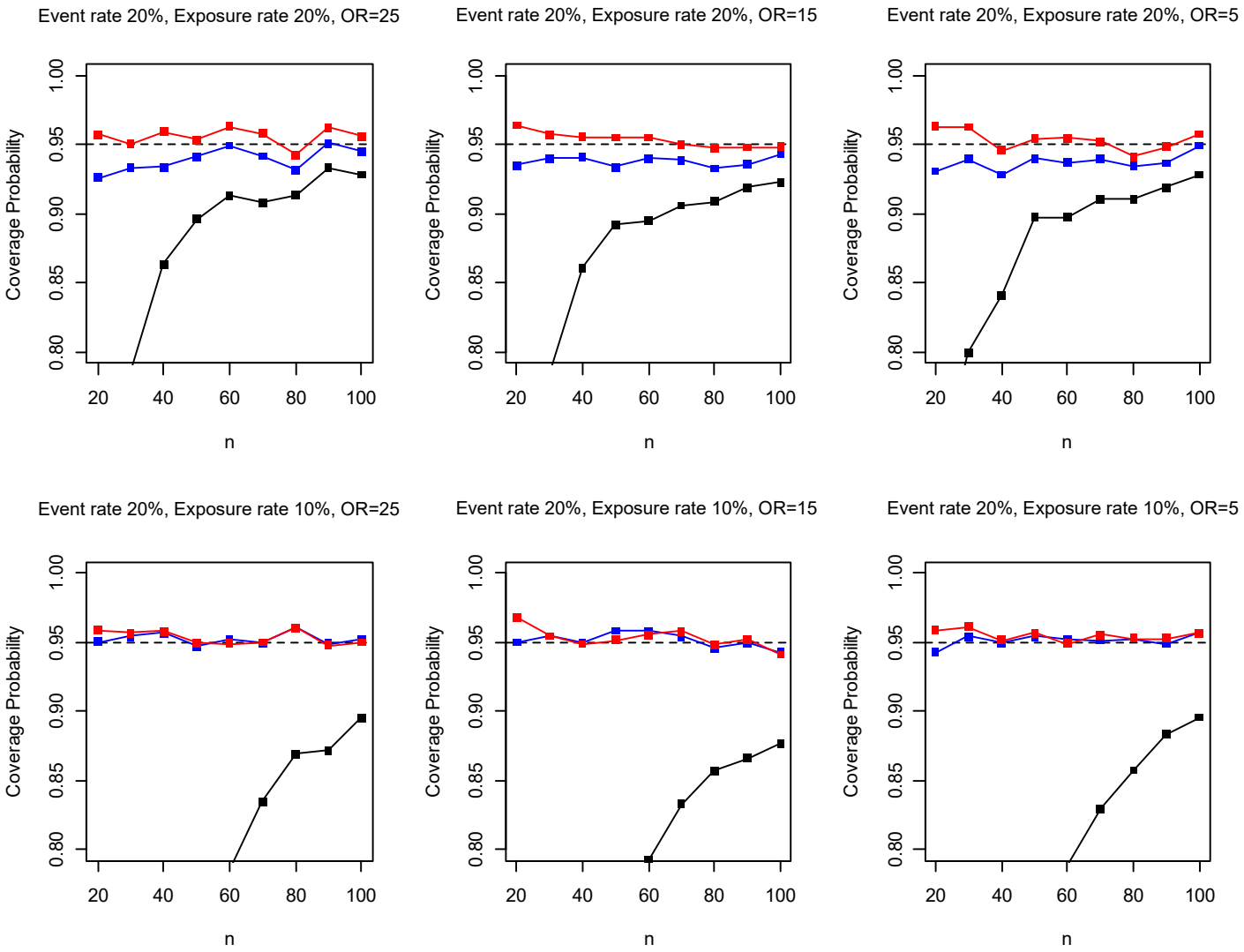

Figure 4. Results of the simulations: modified least-squares regression (**black**: ordinary Wald CI, **blue**: quasi-score CI, **red**: bootstrap CI by the quasi-score statistic; CI: confidence interval).

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1. Results of the modified Poisson and least-squares regression analyses for the epilepsy epidemiologic study ($N = 56$) †.

	Risk ratio estimation by the modified Poisson regression			Risk difference estimation by the modified least-squares regression		
	Quasi-ML estimate (Wald 95%CI)	95%CI by the quasi-score statistic	Bootstrap 95%CI by the quasi-score statistic	Quasi-ML estimate (Wald 95%CI)	95%CI by the quasi-score statistic	Bootstrap 95%CI by the quasi-score statistic
Age at follow up	1.013 (0.982, 1.045)	(0.981, 1.050)	(0.978, 1.067)	0.006 (−0.003, 0.014)	(−0.003, 0.014)	(−0.003, 0.022)
Gender (male vs. female)	1.372 (0.638, 2.950)	(0.669, 3.030)	(0.614, 3.509)	0.097 (−0.089, 0.283)	(−0.063, 0.237)	(−0.071, 0.426)
Mood disorder symptoms	2.076 (0.781, 5.519)	(1.001, 3.230)	(0.882, 7.385)	0.498 (0.153, 0.842)	(0.103, 0.721)	(0.067, 1.000)
Graduating from a school for special needs education	0.240 (0.075, 0.765)	(0.100, 0.686)	(0.028, 0.721)	−0.422 (−0.758, −0.087)	(−0.688, −0.098)	(−0.781, −0.086)

† CI: confidence interval.

medRxiv preprint doi: https://doi.org/10.1101/2025.01.10.25320320; this version posted January 10, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .