Bayesian estimation of the measurement of interactions in epidemiological studies

Shaowei Lin[1,†], Chanchan Hu[1,†], Zhifeng Lin [1], and Zhijian Hu [1,2]*

[1] Department of Epidemiology and Health Statistics, School of Public Health, Fujian Medical University, Fuzhou 350122, China

[2] Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Fujian Medical University, Fuzhou 350122, China

[†] These authors contributed equally to this work.

*Corresponding author: Zhijian Hu, Fujian Medical University, 1 Xue Yuan Road, University Town, Fu Zhou, Fujian 350122, China. Phone: 059183383362; Email: huzhijian@fjmu.edu.cn.

12 **Abstract**

13 **Background** Interaction identification is important in epidemiological studies and can be

14 detected by including a product term in the model. However, as Rothman noted, a product term

15 in exponential models may be regarded as multiplicative rather than additive to better reflect

16 biological interactions. Currently, the additive interaction is largely measured by the relative

17 excess risk due to interaction (RERI), the attributable proportion due to interaction (AP), and the

18 synergy index (S), and confidence intervals are developed via frequentist approaches. However,

19 few studies have focused on the same issue from a Bayesian perspective. The present study aims

20 to provide a Bayesian view of the estimation and credible intervals of the additive interaction

21 measures.

22 **Methods** Bayesian logistic regression was employed, and estimates and credible intervals were

23 calculated from posterior samples of the RERI, AP and S. Since Bayesian inference depends only

24 on posterior samples, it is very easy to apply this method to preventive factors. The validity of the

25 proposed method was verified by comparing the Bayesian method with the delta and bootstrap

26 approaches in simulation studies with example data.

27 **Results** In all the simulation studies, the Bayesian estimates were very close to the

28 corresponding true values. Due to the skewness of the interaction measures, compared with the

29 confidence intervals of the delta method, the credible intervals of the Bayesian approach were

30 more balanced and matched the nominal 95% level. Compared with the bootstrap method, the

31 Bayesian method appeared to be a competitive alternative and fared better when small sample

32 sizes were used.

33 **Conclusions** The proposed Bayesian method is a competitive alternative to other methods. This

34 approach can assist epidemiologists in detecting additive-scale interactions.

36 **Keywords** interaction, additive scale, Bayesian, preventive factor

39   Interaction occurs when the impact of an independent variable (X) on a dependent variable

40   (Y) varies at different levels of a moderating variable (Z)(Andersson, Cuervo-Cazurra & Nielsen,

41   2014). In epidemiology, identifying the interaction between two factors for disease risk is

42   important(Szklo, 2004) because this interaction is significantly associated with disease prevention

43   and intervention. For example, from the perspective of public health, the combined effect of

44   smoking and asbestos on lung cancer surpasses the combination of the individual effects, and a

45   reduction in either factor would also reduce the risk attributable to the other factor in terms of

46   developing lung cancer(Ngamwong et al., 2015).

47   Epidemiologists frequently employ exponential models such as logistic regression and Cox

48   regression to analyze the disease rates and risks(Rothman, Greenland & Lash, 2008). Under an

49   exponential model, there are two scales of interaction: additive and multiplicative. The latter is

50   often employed to assess interactions by including a product interaction term in an exponential

51   model, which implies that the combined effect is larger (or smaller) than the product of the

52   individual effects. However, Rothman and other authors(Rothman, 1976; Andersson et al., 2005;

53   VanderWeele & Robins, 2007) have argued that the examination of interactions on an additive

54   scale makes more sense than that on a multiplicative scale, in which a positive or negative

55   interaction on an additive scale means that the effect in combination is larger or smaller than the

56   sum of the individual effects, respectively.

57   Rothman(Rothman, 1986) proposed different measures to estimate interactions on an additive

58   scale by means of relative risk, such as the relative excess risk due to interaction (RERI), the

59   attributable proportion due to interaction (AP) and the synergy index (S). Hosmer and

60   Lemeshow(Hosmer & Lemeshow, 1992) introduced how to use the delta method to calculate a

61   symmetric confidence interval in a logistic regression model. Furthermore, Knol et al.(Knol et al.,

62   2007) generalized those measures as a departure from additivity in the case of continuous

63   determinant factors. However, due to the skewness of these measures, methods of asymmetric

64   confidence intervals, such as the bootstrap percentile confidence interval(Assmann et al., 1996),

65   variance estimate recovery method(Zou, 2008), and profile likelihood(Richardson & Kaufman,

66   2009), have been developed and have produced competitive results (Andersson et al., 2005; Kuss,

67   Schmidt-Pokrzywniak & Stang, 2010; Nie et al., 2010). In addition to frequentist methods, Chu et

68   al.(Chu, Nie & Cole, 2011) described the use of the Bayesian method to estimate the RERI in a

69   linear additive odds ratio model, which hardly worked with the AP and S.

70   However, for all the methods mentioned above, the estimation of interactions on an additive

71   scale can be applied only to risk factors rather than to preventive factors(Yang et al., 2010;

72   Chatterjee, Shi & García-Closas, 2016; Olsson, Barcellos & Alfredsson, 2017). When either of the

73   two factors at discussion is not a risk factor, both factors should be considered separately or jointly

74   by choosing the low-risk category as the reference level(Rothman, Greenland & Lash, 2008; de

75   Mutsert et al., 2009; Knol et al., 2011). However, this strategy involves applying a logistic model

76   twice. The risk level or combination of factors are first determined, after which the coefficients of

77   the factors are estimated, with the low risk level serving as a reference. Notably, the relation of the

78   estimated coefficients between the two logistic models can be revealed by resetting only the

79   reference level; however, the calculation of confidence intervals becomes complicated with the

80   troublesome involvement of covariance matrices.

81   This article describes how credible intervals of the three measures of interaction on an additive

82   scale can be evaluated using the Bayesian method. Credible intervals are analogous to the

83   confidence intervals in frequentist statistics(Bolstad WM, 2016), although they differ on a

84   philosophical basis(VanderPlas, 2014). The Bayesian method is commonly employed in

85   epidemiology and other applications(Ashby, 2006; MacLehose et al., 2009; Hamra et al., 2013;

86   Ahrens & Pigeot, 2014) due to the simplicity of Bayesian inference. Bayesian inference depends

87   only on posterior samples of the parameters. Once a posterior sample is obtained, a posterior

88   sample of any function of the parameters can be obtained by applying the corresponding function

89   to the posterior sample of those parameters. A sample from the posterior distribution of coefficients

90   of the logistic model can be generated using Markov chain Monte Carlo (MCMC) methods with a

91   random walk Metropolis algorithm(Bolstad WM, 2016). Once a sample is obtained, regardless of

92   whether the logistic model includes preventive factors or risk factors, the posterior samples of the

93 three measures of interaction on an additive scale can be calculated directly by applying the

94 corresponding functions. If there are preventive factors, the functions are composed of two parts:

95 the coefficients corresponding to the new reference (Appendix A) and the measure's functions.

96 Thus, the logistic model requires only one execution. Based on the posterior samples, credible

97 intervals, such as the highest posterior density interval or equal-tailed interval, can easily be

98 assessed.

99

100 **Measures of additive interaction**

101 Consider two binary factors, A and B, with values of 0 and 1, where 0 denotes the absence of

102 a factor and 1 denotes its presence. Four possible combinations exist: A_0B_0, A_1B_0, A_0B_1 and A_1B_1.

103 Suppose the reference group is A_0B_0, and let RR_{10}, RR_{01}, and RR_{11} denote the relative risks for

104 groups A_1B_0, A_0B_1 and A_1B_1, respectively. Rothman(Rothman, Greenland & Lash, 2008)

105 developed three measures of interaction on an additive scale: the relative excess risk due to

106 interaction (RERI),

107 $$RERI = RR_{11} - RR_{10} - RR_{01} + 1,$$

108 the attributable proportion due to interaction (AP),

109 $$AP = \frac{RERI}{RR_{11}} = \frac{RR_{11} - RR_{10} - RR_{01} + 1}{RR_{11}},$$

110 and the synergy index (S)

111 $$S = \frac{RR_{11} - 1}{(RR_{10} - 1) + (RR_{01} - 1)} = \frac{RR_{11} - 1}{RR_{10} + RR_{01} - 2}.$$

112 Note that both the RERI and AP can range from $-\infty$ to $+\infty$, while S can range from 0 to $+\infty$. If no

113 interaction is present on the additive scale, both the RERI and AP will be equal to 0, and S will be

114 equal to 1; if there is more than additivity, both the RERI and AP will be greater than 0, and S will

115 be greater than 1; if there is less than additivity, both the RERI and AP will be less than 0, and S

116 will be less than 1.

117 The RR is readily available in cohort studies. However, in case-control studies, the RR can be

118   approximated by the odds ratio (OR) if the prevalence is sufficiently rare (Appendix B). As

119   suggested by Rothman(Rothman, Greenland & Lash, 2008), the RERI, AP and S can be estimated

120   via multiple logistic regression with indicator variables created for categories A_1B_0, A_0B_1 and

121   A_1B_1. Given that

122    $$logit(p) = \log(\frac{p}{1-p}) = \beta_0 + \beta_1 I(A_1B_0) + \beta_2 I(A_0B_1) + \beta_3 I(A_1B_1) + X\gamma \qquad (1)$$

123   where $p = P(case \mid A, B, X)$ is the corresponding probability of the case given factors A, B and X

124   (X is a vector of the potential confounders) and $I(\cdot)$ is the indicator function; then, $OR_{10} = e^{\beta_1}$,

125    $OR_{01} = e^{\beta_2}$ and $OR_{11} = e^{\beta_3}$ can be substituted for the RRs in the three measures of interaction. As

126   noted by Hosmer and Lemeshow (Hosmer David W, 2013), the ORs can be represented as

127    $OR_{10} = e^{\eta_1}$, $OR_{01} = e^{\eta_2}$ and $OR_{11} = e^{\eta_1 + \eta_2 + \eta_3}$ based on logistic regression with factors A and B and

128   the product of A and B

129    $$logit(p) = \eta_0 + \eta_1 A + \eta_2 B + \eta_3 AB + X\gamma . \qquad (2)$$

130   Because both models are saturated, the two estimations are equivalent.

131

132   **Methods for calculating confidence intervals**

133   The methods used to calculate confidence intervals fall into two categories: symmetric

134   intervals based on a normal distribution and asymmetric intervals based on quantile estimation.

135   Let lnS denote S on the log scale and Z denote any one of the three measures RERI, AP or lnS. As

136   proposed by Hosmer and Lemeshow, the confidence interval of Z is obtained by assuming a normal

137   distribution for Z. The variance in Z can be estimated using the delta method, which employs the

138   first-order approximation of a Taylor series(Hosmer & Lemeshow, 1992). Therefore, the 95%

139   confidence interval of the delta method for Z, which is symmetric about the point estimate, is given

140   by $\hat{Z} \pm 1.96 \times \sqrt{\hat{D}(Z)}$, where $\hat{Z}$ and $\hat{D}(Z)$ are the point estimate and the variance estimate for Z,

141   respectively. The confidence intervals for the RERI and AP can be obtained directly, while the

142 confidence interval for S can be obtained using an exponential function. Zou suggested a recover

143 variance estimate of the variance for measures that considers Taylor series expansion using the

144 multivariate delta method.

145   Another way to estimate confidence intervals is through bootstrapping (Efron, B., &

146 Tibshirani, R.J., 1994). This technique allows estimation of the sampling distribution of almost

147 any statistic using random sampling, as do the measures of additive interaction. Let Z denote any

148 one of the three measures RERI, AP or S. The bootstrap samples are resampled from the original

149 sample, and then Z can be estimated in each of the bootstrap samples. When obtaining the

150 distribution of Z, deriving estimates of standard errors and confidence intervals is straightforward.

151 There are several methods for constructing confidence intervals from the bootstrap distribution.

152   The first of these methods is the normal bootstrap method. The method assumes that Z follows

153 a normal distribution. The variance $\hat{D}(Z)$ is estimated from the bootstrap samples, and the

154 confidence interval is calculated as $\hat{Z} \pm 1.96 \times \sqrt{\hat{D}(Z)}$ Alternatively, the quantile method is the

155 most popular method for constructing confidence intervals. Let Z_{α}^{*} denote the α percentile of the

156 bootstrap-estimated distribution for Z. Efron and Tibshirani (Efron, B., & Tibshirani, R.J., 1994)

157 used percentiles of the bootstrap distribution to estimate the $1-\alpha$ confidence interval for Z as

158 (percentile bootstrap, PB)

159              $(Z_{\alpha/2}^{*}, Z_{1-\alpha/2}^{*})$.

160 Davison and Hinkley (A. C. Davison, 1997) proceeded in a similar way, but a different formula

161 was used to construct the confidence interval for Z, as (basic bootstrap, BB)

162              $(2\hat{Z} - Z_{1-\alpha/2}^{*}, 2\hat{Z} - Z_{\alpha/2}^{*})$.

163

164 **Bayesian estimation of the additive measure**

165   Given a model, there are three steps involved in Bayesian analysis(Bolstad WM, 2016): (1)

166 determining the marginal likelihood of the data, (2) specifying prior probabilities for the

167 parameters, and (3) applying Bayes' theorem to predictor variables and the outcome variable via

168   Bayesian modeling.

169   Let Y denote the binary outcome variable. Consider the logistic regression of the form

170                    $$Y: \; Bernoulli(p)$$

171   where $Bernoulli(\cdot)$ is a Bernoulli distribution and p satisfies Eq. (1) or Eq. (2). The likelihood

172   for the logistic regression is

173             $$p(x \mid \theta) = \prod_i h(x;\theta)^Y (1 - h(x;\theta))^{1-Y}$$

174   where θ is a parameter (such as $\theta = (\beta_0, \beta_1, \beta_2, \beta_3, \gamma)$ in Eq. (1)) and $h(x;\theta) = \dfrac{1}{1+e^{-logit(p)}}$ is the

175   predicted probability that Y is 1.

176   The prior probability is an unconditional probability that is assigned before any data are

177   accounted for. There are three main categories of prior distributions: informative priors, weakly

178   informative priors and noninformative priors. When a family of conjugate priors exists, the

179   conjugate priors are chosen for computational efficiency. However, there is no conjugate prior for

180   the likelihood function in logistic regression. When Bayesian inference is performed analytically,

181   this makes the posterior distribution difficult to calculate except in very low dimensions. However,

182   software such as OpenBUGS(Lunn et al., 2009), JAGS(Plummer, 2003) and Stan(Carpenter et al.,

183   2017) allows these posteriors to be computed via simulation; hence, a lack of conjugacy is not a

184   concern. This paper employs noninformative or weakly informative priors $p(\theta)$ for all model

185   parameters.

186   According to Bayes' theorem, the joint posterior distribution of the model parameters is

187   proportional to the product of the likelihood and priors,

188        $$p(\theta \mid x) \propto p(\theta) \cdot p(x \mid \theta) = p(\theta) \prod_i h(x;\theta)^Y (1 - h(x;\theta))^{1-Y}.$$

189   This equation is too complex to solve analytically; therefore, Monte Carlo methods are often used

190   to summarize the posterior distribution. A combination of Monte Carlo and Markov chain can

191   result in effective sampling and evaluation. Thus, this posterior computation can be accomplished

192   easily using Markov chain Monte Carlo (MCMC) methods with a random walk Metropolis

193   algorithm. Since RERI, AP and S are functions of θ, the values of the three measures of interaction

194   can be calculated directly when a sample is generated from the posterior distribution $p(\theta\,|\,x)$.

195   The uncertainties of the RERI, AP and S can be summarized by giving a range of values on

196   the posterior probability distribution that includes 95% of the probability, $(P_{2.5}, P_{97.5})$, which is

197   called a 95% credibility interval (calculated based on an equal-tailed interval). Similarly, in a

198   Bayesian credibility interval, the null hypothesis is rejected if the credibility interval does not

199   encompass the parameters of the null hypothesis. For example, there is an interaction if the

200   credibility interval of an RERI does not contain 0. Notably, the confidence interval captures the

201   uncertainty about the interval obtained (whether it contains the true value or not), and it cannot be

202   interpreted as a probabilistic statement about the true parameter value. However, the credible

203   interval captures the current uncertainty in the location of the parameter value and thus can be

204   interpreted as a probabilistic statement about the parameter.

205

206   **Preventive factors**

207   Measures of interaction were developed for risk factors, so it is not appropriate to use them directly

208   when a preventive factor exists. The recoded method, which can be seen as a choice of reference

209   category, is used to address this issue(Knol et al., 2011). The use of the recoded method and the

210   use of an updated model is a cumbersome and tedious process. In fact, when the reference category

211   changes, the new logistic regression coefficients $\theta' = (\beta_0', \beta_1', \beta_2', \beta_3')$ are completely dependent on

212   the original logistic coefficients $\theta = (\beta_0, \beta_1, \beta_2, \beta_3)$, as resetting the reference category is equivalent

213   to subtracting the new reference category coefficient from θ (Hosmer David W, 2013). The

214   relationships between θ' and θ for the three interaction measures are shown in Appendix A. In the

215   Bayesian approach, the three measures of interaction are calculated directly by setting the

216   reference category as that whose coefficient is the smallest (set $\beta_0 = 0$ in the dummy coding

217   scheme). Thus, it is not necessary to update the model, and the uncertainty of the parameters will

218   be obtained from the posterior sample. The R(R Development Core Team, 2023, https://www.r-

219   project.org/) function for computing new coefficients based on the lowest risk category as a

220   reference category is provided in Appendix C.

221

Simulation study setting

223   To assess the performance of Bayesian estimation of the additive measure, simulation studies

224   were conducted to compare the results of confidence interval estimation by these methods.

225   Regardless of the effect measure used, OR, RR, or hazard ratio, the evaluation process was similar,

226   so a case–control design was employed. The values of OR_{01}, OR_{10} and OR_{11} were set as in the study

227   by Assmann (Assmann et al., 1996), where 20 scenarios ($2\,OR_{01} \times 2\,OR_{10} \times 5\,OR_{11}$) are described

228   in Table 1. These scenarios include situations with strong synergy, weak synergy, no interaction,

229   weak antagonism or strong antagonism.

230   In each scenario, the sample size was set to n=600, and the factor exposure rate was fixed, in

231   which the proportions of subjects exposed to factor A alone, factor B alone, and both factors A

232   and B were 0.1, 0.2 and 0.1, respectively. In addition to the parameters mentioned above, two

233   additional parameters were considered: the proportion of the case number relative to the sample

234   size, which changes from $p_1 = 1/2$ to $p_2 = 1/3$, i.e., the number of cases is reduced from 300 to

235   200; and the other is where there are potential confounders X and the coefficient of the variable is

236   set to $\gamma = 0.5$ when containing a confounding variable.

237   Three simulation studies were designed to evaluate the effects of the methods: balanced design

238   ($p_1 = 1/2$), unbalanced design ($p_2 = 1/3$) and balanced design with a confounding variable. For

239   each combination of parameters in each study, 1000 replicates were performed. Bayesian

240   estimation was compared with the delta method and bootstrap method. As suggested by Assmann,

241   due to the skewed distribution of the three measures of interaction, the percentile bootstrap was

242   suitable. Thus, only the percentile bootstrap was performed in the simulation studies. The bootstrap

243   sampling, with 1000 resamples, was performed separately within cases and controls. Therefore,

244   the number of cases and controls in the bootstrap samples were identical to those in the original

245   sample. Therefore, for each sample, point estimates of the interaction measures were calculated,

246 as were 95% confidence intervals for the delta method and percentile bootstrap method and 95%

247 credible intervals for the Bayesian method.

248   The posterior distributions of the parameters were estimated via the MCMC method with a

249 random walk Metropolis algorithm, which can be performed easily via the R package

250 MCMCpack(Martin, Quinn & Park, 2011). Specifically, the Markov chain was run for a total of

251 20000 iterations with 10000 burn-in iterations. To monitor convergence, multiple chains with

252 different initial values were constructed, and trace plots were examined visually. The convergence

253 of Markov chains was further assessed using the standard convergence statistic. Because an

254 improper prior distribution may lead to inaccurate posterior estimates, proper but diffuse prior

255 distributions with weak information for the parameters were employed(Spiegelhalter, 1995).

256 Alternatively, noninformative priors, such as the Jeffreys prior or uniform prior with proper

257 bounds, may also be used(Bolstad WM, 2016). Specifically, in all the simulation studies, the prior

258 distributions of the parameters were set to a normal distribution with a large variance,

$$\theta = (\beta_0, \beta_1, \beta_2, \beta_3) \sim N(0, 10^2 I_4)$$

259

$$\gamma \sim N(0, 10^2)$$

260

261 where I_4 is the identity matrix of size 4, which is a 4×4 square matrix with ones on the main

262 diagonal and zeros elsewhere.

263   Three criteria were used to evaluate the extent to which the empirical coverage of the

264 confidence interval matched the nominal 95% level. The first was the rank of the deviation from

265 the nominal level, which was obtained by ranking the absolute value of the difference between the

266 coverage rate and 95%. The second was the percentage of times that the interval covered the true

267 value, which fell within the range of 93.6-96.4% (due to the 1000 samples). The third was the

268 percentage of missed data in either direction. The left miscoverage rate was the proportion of times

269 the lower limits were larger than the true value, while the right miscoverage rate was the proportion

270 of times the upper limits were less than the true value. A balanced miscoverage indicated that the

271 confidence interval was neither too wide nor too narrow. From the usual sample-size formula for

272   estimating proportions, if the true coverage rate for a particular type of confidence interval (or

273   credible interval) was 95%, then the probability of obtaining coverage between 92.5% and 97.5%

274   would be approximately 0.95. Thus, an advertised 95% confidence interval (or credible interval)

275   should miss approximately 2.5% from each side; i.e., the left/right miscoverage rates should be

276   2.5%.

277

278   **Results of the simulation study**

279   With respect to the balanced design with 300 cases and controls, for each scenario and each

280   method, there were 1000 point estimates of the interaction measures with considerable variation.

281   However, for each interaction measure, the mean and median of the point estimates were very

282   close to the corresponding true values. For instance, in scenario A1, the true value of the RERI

283   was 12. The 1000-point estimates of delta for the RERI ranged from -2.83 to 32.91; the mean and

284   median values were 13.21 and 11.95, respectively, and the standard deviation was 7.62.

285   The sampling distributions of those point estimates, estimated from the same datasets, were

286   almost the same among the three methods. This was consistent with expectations. Figure 1 shows

287   the sampling distribution of each interaction measure in scenarios A1, A3 and A5, corresponding

288   to strong synergy, no interaction, and strong antagonism, respectively. The RERI seemed to follow

289   a normal distribution with a mean of 0 when there was no interaction but a mildly right-skewed

290   distribution when there was synergy and a mildly left-skewed distribution when there was

291   antagonism. For AP, the distributions were skewed to the left in all scenarios except for

292   antagonism. For S, the distributions were skewed to the right in every scenario.

293   Table 2 summarizes the coverage performance of the three estimation methods for the RERI.

294   Due to the skewed distribution, when using the median as the estimate, all the values were close

295   to the true value of the RERI. For the delta, bootstrap, and Bayesian intervals, the average ranks

296   of the deviation were 2.75, 1.50 and 1.65, respectively. This shows that, over a wide range of

297   scenarios, the bootstrap method most often yields a coverage rate closest to 95%, followed by the

298   Bayesian method. Notably, the rank would be equal to 3 if the coverage rate was the farthest from

299   95% in every scenario, which indicates that the delta intervals least accurately match the nominal

300   value of 95%.

301   For delta intervals, the coverage rate fit into the target range of 93.6-96.4% in 4 of the 20

302   scenarios. Clearly, the coverage rate tended to be low when there was synergy and high when there

303   was antagonism. In particular, in scenarios with synergy, delta intervals covered the true value of

304   the RERI by lying too far to the right, which resulted in low right miscoverage and high left

305   miscoverage. However, in scenarios with antagonism, the delta intervals may be too wide,

306   resulting in a high coverage rate and low miscoverage. In contrast, the Bayesian coverage was

307   within the target range in 18 of the 20 scenarios. Furthermore, the left/right miscoverages were

308   approximately 2.5%, which indicates that the Bayesian intervals were more evenly balanced. The

309   bootstrap intervals were similar to those of the Bayesian method.

310   Table 3 summarizes the coverage performance of the three estimation methods for AP. For

311   the delta, bootstrap, and Bayesian intervals, the average ranks of the deviation were 3, 1.45 and

312   1.4, respectively. The coverage for delta intervals was in the target range for only 2 of the 20

313   scenarios, while that for the bootstrap and Bayesian intervals was in the target range for 19 of the

314   20 scenarios. For the delta interval, the miscoverage was lop-sided, with high right miscoverage.

315   It tended to achieve an excessively high lower confidence limit, which would probably lead to

316   false positive results. The Bayesian interval had better coverage, but the miscoverage appeared to

317   be slightly unbalanced, as in the case of the bootstrap interval. Nevertheless, from a practical

318   standpoint, this miscoverage is reasonable.

319   Table 4 summarizes the coverage performance of the three estimation methods for S. For the

320   delta, bootstrap, and Bayesian intervals, the average ranks of the deviation were 2.35, 1.60 and

321   1.85, respectively, and the numbers of points within the target range were 15, 19 and 18,

322   respectively. Notably, the imbalance in miscoverage was slightly alleviated for the three estimation

323   methods.

324   Further simulation results with imbalanced designs or confounding variables are reported in

325   the supplementary material (Table S1-S6) for the interaction measures RERI, AP and S. In general,

326 the performances of both studies exhibited similar statistical behavior as that of the balanced

327 design. Specifically, for AP, the coverage rates of the delta interval were especially low, and the

328 worst coverage rate was as low as 87.6%. However, the coverage rates of the bootstrap and

329 Bayesian intervals were mostly slightly less than 95%. In addition, the bootstrap and Bayesian

330 intervals achieved balanced miscoverage, but the delta interval did not. In summary, the

331 simulations demonstrate that, for three interaction measures, the bootstrap and Bayesian methods

332 performed satisfactorily, while the delta method did not. The difference between the bootstrap and

333 Bayesian methods was minimal and can be considered negligible.

334

335 **Example**

336   This example came from a case–control study among male veterans younger than 60 years

337 concerning smoking and alcohol use in relation to oral cancer(Rothman & Keller, 1972). The data

338 are presented in Table 5. Several authors have used various methods to calculate confidence

339 intervals (CIs) for the RERI. Specifically, based on Hosmer and Lemeshow's delta

340 method(Hosmer & Lemeshow, 1992), the point estimate of the RERI was 3.74, and the asymptotic

341 95% confidence interval was (-1.85 to 9.33). Based on Zou's variance estimates recovery method

342 (MOVER)(Zou, 2008), the 95% confidence interval was (-11.41 to 21.84). Based on Richardson

343 and Kaufman's profile likelihood method(Richardson & Kaufman, 2009), the 95% confidence

344 interval was (-3.29 to 17.21). The 95% confidence interval of the percentile bootstrap method

345 proposed by Assmann et al was (-10.77 to 1.6×10^7)(Assmann et al., 1996), which was too wide

346 due to the sparse cells in some bootstrapped samples. In addition, Chu et al.(Chu, Nie & Cole,

347 2011) proposed a Bayesian method for linear additive odds ratio models and reported that the

348 posterior medians of the RERI with 95% high probability density and credible intervals were 3.00

349 (-2.27 to 8.90), 4.05 (-2.54 to 12.65) and 4.75 (-2.63 to 15.60), corresponding to different prior

350 distributions.

351   For the data at hand, using the proposed Bayesian method, the point estimates with 95%

352 credible intervals for REIR, AP and S are shown in Table 6, along with the confidence intervals

353   of the delta and bootstrap methods. The three interaction measures of all methods suggested

354   synergy, but the 95% intervals were broad and not significant. For the RERI, the Bayesian method

355   provided a reasonable 95% credible interval (-3.93 to 19.59). For AP, owing to the left-skewed

356   distribution, the Bayesian interval was positioned further left than the delta interval was. However,

357   for S, owing to the right-skewed distribution, the Bayesian interval was positioned further right

358   than the delta interval was. Figure 2 shows the distributions of interactions in this example.

359   Furthermore, the features that appeared in the simulation studies seemed to be maintained.

360   Notably, the overall performance of the bootstrap method was comparable to that of the

361   Bayesian method in all the simulation studies. However, in this example, regardless of the three

362   interaction measures, the intervals in the Bayesian method were narrower than those in the

363   bootstrap method, which may indicate that there is less variance in the Bayesian method. Notably,

364   the upper limits of the intervals obtained via the bootstrap method were greater than those obtained

365   via the Bayesian method, especially for the RERI. This indicates that the percentile bootstrap may

366   perform poorly when data with a small sample size exist in some groups, but the proposed

367   Bayesian approach is free from this predicament.

368

369   **Discussion**

370   It is very important to assess whether an additive interaction exists between two factors during

371   the interpretation of epidemiologic data[4] because this information is relevant to disease prevention

372   and intervention. In general, in addition to hypothesis tests, confidence intervals are used in the

373   assessment of statistical interactions(Greenland et al., 2016). There are many methods for

374   constructing confidence intervals for RERI, such as the delta method(Hosmer & Lemeshow,

375   1992), bootstrap method(Assmann et al., 1996), variance estimate recovery approach(Zou, 2008),

376   and profile likelihood approach(Richardson & Kaufman, 2009), based on the theories of

377   classical frequentists. However, due to the complexity of the formulas, there is little information

378   available for AP or S. An alternative method is the Bayesian approach, which uses a credible

379   interval rather than a confidence interval. Chu et al.(Chu, Nie & Cole, 2011) proposed a Bayesian

380 method for linear additive odds ratio models to construct the credible interval of the RERI but not

381 that of the AP or S. This article proposes a Bayesian method to estimate all the additive interaction

382 measures, RERI, AP and S, along with their 95% credible intervals. In contrast to Chu's method,

383 in which the posterior distribution is a step function, this method does not require inequality

384 constraints on the parameters, which may expand the sample space of the parameter and lead to

385 higher efficiency when MCMC is applied.

386   The simulation studies show that the distributions of the additive interaction measures are

387 skewed, especially those of AP and S. Therefore, the delta method, whose Wald-type confidence

388 interval is symmetric, performs poorly in terms of the empirical coverage of the confidence interval

389 or the percentage of times missed in either direction. By recovering the variances, the method

390 proposed by Zou may alleviate this situation. However, as Wald-type confidence intervals are

391 based on asymptotic variance, they may be misleading in nonmultiplicative models(Moolgavkar

392 & Venzon, 1987).

393   Compared with those of the delta method, the Bayesian credible intervals proposed here and

394 the bootstrap confidence intervals, which do not rely on asymptotic approximations, are more

395 likely to catch on each side of the true value, and their coverage is close to the nominal level of

396 95%. The bootstrap confidence intervals, as proposed by Assmann et al., are preferred in practice

397 (Töpper et al., 2018). However, the bootstrap method has several limitations. DiCiccio and

398 Efron(DiCiccio & Efron) cautioned that a bootstrap sample may depend on a distribution with

399 difficult computations in nonmultiplicative models. In addition, as observed in the example, the

400 confidence interval could come to nothing when the sample is insufficient(Chernick, 1999).

401 However, Bayesian credible intervals have minimal impact and work well.

402   Some authors(Rothman, Greenland & Lash, 2008; Knol et al., 2011) have noted that additive

403 measures are suitable only for risk factors. If there is a preventive factor, it should be recoded to a

404 risk factor before calculating these measures. To determine where it is a preventive factor and how

405 to recode it, the logistic regression model needs to be applied at least twice. When this strategy is

406 applied to the bootstrap method, each bootstrap sample needs to be recoded separately rather than

407 used for the bootstrap samples. This approach significantly increases the calculation time.

408 However, in the Bayesian framework, once a sample of the posterior distribution is obtained, one

409 can apply this sample to any function(Bolstad WM, 2016). Therefore, the posterior distributions

410 of the additive measures can be estimated when the coefficients $\theta = (\beta_0, \beta_1, \beta_2, \beta_3)$ are obtained

411 by the Bayesian method. This makes the model presented here highly flexible. As shown in

412 Appendix A, the additive measures are computed directly after parameter transformation, but the

413 computation time can be ignored. Moreover, one can adjust the confounding variable X in various

414 forms in logistic regression, such as the linear or spline function of X.

415   The frequentist and Bayesian methods appear to be superficially opposite, but they can be

416 considered complementary in practice(Chu, Nie & Cole, 2011). When using a weak prior

417 distribution in Bayesian methods, the inference is often consistent with that of frequentists.

418 However, the Bayesian method is attractive because proper prior distributions can be determined

419 from past information, such as subject-matter knowledge or previous experiments(Bolstad WM,

420 2016). Furthermore, the credible interval is constructed directly from the highest probability

421 density of the posterior samples, and it can capture the uncertainty of the parameter values without

422 having to rely on asymptotic approximations. This approach can often yield better results,

423 especially for small samples, as shown in the example. In addition, Bayesian methods have a range

424 of other advantages, including convenient application to other indicators related to interactions

425 (e.g., PRISM(Lee, 2013; Lin & Lee, 2016), CPWs(Lee & Wu, 2023)) and the ability to explore

426 disease attribution to multiple exposures and their interactions.

427   In this paper, a Bayesian method for the estimation of additive measures is proposed. By

428 rearranging the posterior sample, it works for both risk factors and preventive factors. The

429 simulation studies and example results show that the Bayesian method is a competitive alternative

430 to other methods.

431

432 **Appendix A**

433 **Computation of the additive measure in the logistic model based on the new reference**

434   **category**

435   Considering factors A and B jointly, four potential categories exist: A_0B_0, A_1B_0, A_0B_1 and

436   A_1B_1. To simplify the description, using the overparameterized dummy coding scheme, the

437   coefficients are β_0, β_1, β_2 and β_3 in logistic regression. When using A_0B_0 as the reference

438   category ($\beta_0 = 0$ in the dummy coding scheme), the ORs can be rewritten as $OR_{00} = e^{\beta_0 - \beta_0}$,

439   $OR_{10} = e^{\beta_1 - \beta_0}$, $OR_{01} = e^{\beta_2 - \beta_0}$ and $OR_{11} = e^{\beta_3 - \beta_0}$. The three measures of interaction can be

440   equivalently written as follows

$$RERI = OR_{11} - OR_{10} - OR_{01} + OR_{00} = e^{-\beta_0} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}),$$

$$AP = e^{-\beta_0} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}) / e^{\beta_3 - \beta_0} = e^{-\beta_3} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}),$$

$$S = \frac{e^{\beta_3} - e^{\beta_0}}{e^{\beta_1} + e^{\beta_2} - 2e^{\beta_0}}.$$

444   When choosing the category with the lowest risk when both factors are considered jointly as the

445   reference category, for example, group A_1B_0, the coefficients β' s of the four categories can be

446   expressed as $\beta_0' = \beta_0 - \beta_1$, $\beta_1' = \beta_1 - \beta_1$, $\beta_2' = \beta_2 - \beta_1$ and $\beta_3' = \beta_3 - \beta_1$. Then, the ORs can be

447   expressed as $OR_{00}' = e^{\beta_0 - \beta_1}$, $OR_{10}' = e^{\beta_1 - \beta_1}$, $OR_{01}' = e^{\beta_2 - \beta_1}$ and $OR_{11}' = e^{\beta_3 - \beta_1}$. The three measures of

448   interaction can be computed by rearranging the ORs, which have the form of

$$RERI = OR_{01}' - OR_{00}' - OR_{11}' + OR_{10}' = -e^{-\beta_1} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}),$$

$$AP = RERI / OR_{01}' = -e^{-\beta_2} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}),$$

$$S = \frac{e^{\beta_2} - e^{\beta_1}}{e^{\beta_3} + e^{\beta_0} - 2e^{\beta_1}}.$$

452   Similar results occur when the reference category is set as group A_0B_1or A_1B_1. When the reference

453   category is A_0B_1, the three measures of interaction have the form of

$$RERI = OR_{10}' - OR_{11}' - OR_{00}' + OR_{01}' = -e^{-\beta_2} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}),$$

455
$$AP = RERI / OR'_{10} = -e^{-\beta_1} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}),$$

456
$$S = \frac{e^{\beta_1} - e^{\beta_2}}{e^{\beta_3} + e^{\beta_0} - 2e^{\beta_2}}.$$

457 When the reference category is A_1B_1, the three measures of interaction have the form of

458
$$RERI = OR'_{00} - OR'_{01} - OR'_{10} + OR'_{11} = e^{-\beta_3} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}),$$

459
$$AP = RERI / OR'_{00} = e^{-\beta_0} \cdot (e^{\beta_3} - e^{\beta_1} - e^{\beta_2} + e^{\beta_0}),$$

460
$$S = \frac{e^{\beta_0} - e^{\beta_3}}{e^{\beta_1} + e^{\beta_2} - 2e^{\beta_3}}.$$

461

462 **Appendix B**

463 ***RR* can be expressed as a function of both the *OR* and the intercept b_0.**

464   If there are independent variable X and dependent variable Y in the logistic regression, then

465 we have:

466
$$\log it \; P(Y = 1) = b_0 + b_1 X \,;$$

467 The probability of an event occurring in the control group ($X=0$) is:

468
$$P_0 = P(Y = 1 | X = 0) = \frac{e^{b_0}}{1 + e^{b_0}};$$

469 The probability of an event occurring in the treatment group ($X=1$) is:

470
$$P_1 = P(Y = 1 | X = 1) = \frac{e^{b_0 + b_1}}{1 + e^{b_0 + b_1}};$$

471 The *OR* can be expressed as:

472
$$OR = \frac{P_1 / (1 - P_1)}{P_0 / (1 - P_0)} = \frac{e^{b_0 + b_1}}{e^{b_0}} = e^{b_1};$$

473 The *RR* can be expressed as:

474
$$RR = \frac{P_1}{P_0} = \frac{e^{b_1}(1 + e^{b_0})}{1 + e^{b_0 + b_1}} = \frac{OR \cdot (1 + e^{b_0})}{1 + OR \cdot e^{b_0}} = \frac{OR}{1 - P_0 + P_0 \cdot OR}.$$

Appendix C

R function for computing new coefficients based on the lowest risk category as reference category

```
newCoefficents <-   function(coefficients){          # coefficients: the original coefficients
    if(length(coefficients)!=4) stop("the length of coefficients must be equate to 4")
    minloc <- which.min(coefficients)
    loc <- 1:4
    if(minloc==2) loc<-c(2,1,4,3)
    if(minloc==3) loc<-c(3,4,1,2)
    if(minloc==4) loc<-c(4,3,2,1)
    new.coef<- coefficients[loc]
    new.coef<- new.coef – new.coef[1]
    return(new.coef)
}
```

Reference

Ahrens W, Pigeot I (eds.). 2014. *Handbook of Epidemiology*. New York, NY: Springer New York. DOI: 10.1007/978-0-387-09834-0.

Andersson T, Alfredsson L, Källberg H, Zdravkovic S, Ahlbom A. 2005. Calculating measures of biological interaction. *European Journal of Epidemiology* 20:575–579. DOI: 10.1007/s10654-005-7835-x.

Andersson U, Cuervo-Cazurra A, Nielsen BB. 2014. From the Editors: Explaining interaction effects within and across levels of analysis. *Journal of International Business Studies* 45:1063–1071. DOI: 10/f22ng5.

Applied Logistic Regression | Wiley Series in Probability and Statistics. *Available at*

500   *https://onlinelibrary.wiley.com/doi/book/10.1002/9781118548387*   (accessed   September

501   26, 2023).

502   Ashby D. 2006. Bayesian statistics in medicine: a 25 year review. *Statistics in Medicine* 25:3589–

503   3631. DOI: 10/ffkn7j.

504   Assmann SF, Hosmer DW, Lemeshow S, Mundt KA. 1996. Confidence intervals for measures of

505   interaction. *Epidemiology (Cambridge, Mass.)* 7:286–290. DOI: 10/cn3xm5.

506   Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li

507   P, Riddell A. 2017. Stan: A Probabilistic Programming Language. *Journal of Statistical*

508   *Software* 76:1–32. DOI: 10.18637/jss.v076.i01.

509   Chatterjee N, Shi J, García-Closas M. 2016. Developing and evaluating polygenic risk prediction

510   models for stratified disease prevention. *Nature Reviews. Genetics* 17:392–406. DOI:

511   10.1038/nrg.2016.27.

512   Chernick MR. 1999. *Bootstrap Methods: A Practitioner's Guide*. New York Weinheim: Wiley-

513   Interscience.

514   Chu H, Nie L, Cole SR. 2011. Estimating the relative excess risk due to interaction: a bayesian

515   approach. *Epidemiology (Cambridge, Mass.)* 22:242–248. DOI: 10/fgzpk7.

516   DiCiccio TJ, Efron B. Bootstrap Confidence Intervals. DOI: 10/cmj7pz.

517   Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, Altman DG. 2016.

518   Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations.

519   *European Journal of Epidemiology* 31:337–350. DOI: 10.1007/s10654-016-0149-3.

520   Hamra G, Richardson D, Maclehose R, Wing S. 2013. Integrating informative priors from

521   experimental research with Bayesian methods: an example from radiation epidemiology.

522   *Epidemiology (Cambridge, Mass.)* 24:90–95. DOI: 10/gs6kbw.

523   Hosmer DW, Lemeshow S. 1992. Confidence Interval Estimation of Interaction. *Epidemiology*

524   3:452–456. DOI: 10/fqxb8k.

525   Introduction to Bayesian Statistics, 3rd Edition | Wiley. *Available at*

526   *https://www.wiley.com/enus/Introduction+to+Bayesian+Statistics%2C+3rd+Edition-p-*

527   *9781118593226* (accessed September 26, 2023).

528   Knol MJ, van der Tweel I, Grobbee DE, Numans ME, Geerlings MI. 2007. Estimating interaction

529   on an additive scale between continuous determinants in a logistic regression model.

530   *International Journal of Epidemiology* 36:1111–1118. DOI: 10/cnzcrw.

531   Knol MJ, VanderWeele TJ, Groenwold RHH, Klungel OH, Rovers MM, Grobbee DE. 2011.

532   Estimating measures of interaction on an additive scale for preventive exposures. *European*

533   *Journal of Epidemiology* 26:433–438. DOI: 10/fdkwzr.

534   Lee W-C. 2013. Assessing causal mechanistic interactions: a peril ratio index of synergy based on

535   multiplicativity. *PloS One* 8:e67424. DOI: 10/f49xg3.

536   Lee W-C, Wu Y-C. 2023. Disease Attribution to Multiple Exposures Using Aggregate Data.

537   *Journal of Epidemiology* 33:405–409. DOI: 10.2188/jea.JE20210084.

538   Lin J-H, Lee W-C. 2016. Complementary Log Regression for Sufficient-Cause Modeling of

539   Epidemiologic Data. *Scientific Reports* 6:39023. DOI: 10.1038/srep39023.

540   Lunn D, Spiegelhalter D, Thomas A, Best N. 2009. The BUGS project: Evolution, critique and

541   future directions. *Statistics in Medicine* 28:3049–3067. DOI: 10/bp63vd.

542   MacLehose RF, Olshan AF, Herring AH, Honein MA, Shaw GM, Romitti PA, National Birth

543   Defects Prevention Study. 2009. Bayesian methods for correcting misclassification: an

544   example from birth defects epidemiology. *Epidemiology (Cambridge, Mass.)* 20:27–35.

545   DOI: 10/fnz67z.

546   Martin AD, Quinn KM, Park JH. 2011. MCMCpack : Markov Chain Monte Carlo in *R. Journal*

547   *of Statistical Software*. 42(9): 1-21. DOI: 10.18637/jss.v042.i09.

548   Moolgavkar SH, Venzon DJ. 1987. General relative risk regression models for epidemiologic

549   studies. *American Journal of Epidemiology* 126:949–961. DOI:

550   10.1093/oxfordjournals.aje.a114733.

551   de Mutsert R, Jager KJ, Zoccali C, Dekker FW. 2009. The effect of joint exposures: examining

552   the presence of interaction. *Kidney International* 75:677–681. DOI: 10/fsmkrn.

553   Ngamwong Y, Tangamornsuksan W, Lohitnavy O, Chaiyakunapruk N, Scholfield CN, Reisfeld

554   B, Lohitnavy M. 2015. Additive Synergism between Asbestos and Smoking in Lung

555   Cancer Risk: A Systematic Review and Meta-Analysis. *PloS One* 10:e0135798. DOI:

556   10/gmwr58.

557   Olsson T, Barcellos LF, Alfredsson L. 2017. Interactions between genetic, lifestyle and

558   environmental risk factors for multiple sclerosis. *Nature Reviews Neurology* 13:25–36.

559   DOI: 10/f3vhw6.

560   Plummer M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs

561   sampling. Proceedings of the 3rd International Workshop on Distributed Statistical

562   Computing (DSC 2003), Vienna, 20-22 March 2003, 1-10.

563   R Core Team (2023). R: A language and environment for statistical computing. R Foundation for

564   Statistical Computing, Vienna, Austria. URL https://www.r-project.org/.

565   Richardson DB, Kaufman JS. 2009. Estimation of the relative excess risk due to interaction and

566   associated confidence bounds. *American Journal of Epidemiology* 169:756–760. DOI:

567   10.1093/aje/kwn411.

568   Rothman KJ. 1976. THE ESTIMATION OF SYNERGY OR ANTAGONISM1. *American*

569   *Journal of Epidemiology* 103:506–511. DOI: 10/gs6j69.

570   Rothman KJ. 1986. *Modern epidemiology*. Boston : Little, Brown.

571   Rothman K, Greenland S, Lash TL. 2008. Modern Epidemiology, 3rd Edition.

572   Rothman K, Keller A. 1972. The effect of joint exposure to alcohol and tobacco on risk of cancer

573   of the mouth and pharynx. *Journal of Chronic Diseases* 25:711–716. DOI: 10/cdxfz2.

574   Spiegelhalter WRG S Richardson, David (ed.). 1995. *Markov Chain Monte Carlo in Practice*.

575   New York: Chapman and Hall/CRC. DOI: 10.1201/b14835.

576   Szklo M (Moyses). 2004. *Epidemiology : beyond the basics*. Sudbury, Mass. : Jones and Bartlett.

577   VanderPlas J. 2014. Frequentism and Bayesianism: A Python-driven Primer. In: Austin, Texas,

578   85–93. DOI: 10/gmcdmn.

579   VanderWeele TJ, Robins JM. 2007. The Identification of Synergism in the Sufficient-Component-

580   Cause Framework. *Epidemiology* 18:329–339. DOI: 10/cmq9qc.

581   Yang X, Ma RCW, So WY, Ko GTC, Kong APS, Zhao H, Xu G, Tong PCY, Chan JCN. 2010.

582   White blood cell count and renin–angiotensin system inhibitors for the risk of cancer in

583   type 2 diabetes. *Diabetes Research and Clinical Practice* 87:117–125. DOI: 10/bfnh2d.

584   Zou GY. 2008. On the Estimation of Additive Interaction by Use of the Four-by-two Table and

585   Beyond. *American Journal of Epidemiology* 168:212–224. DOI: 10/b4nc3w.

586

Figure 1

Figure 1. Sampling distributions of interaction measures, which were estimated from the scenarios A1 (top), A3 (middle) and A5 (bottom) in the balanced design simulation study.

The lines reveal the median of the distribution which were very close to the true value of RERI, AP and S respectively.

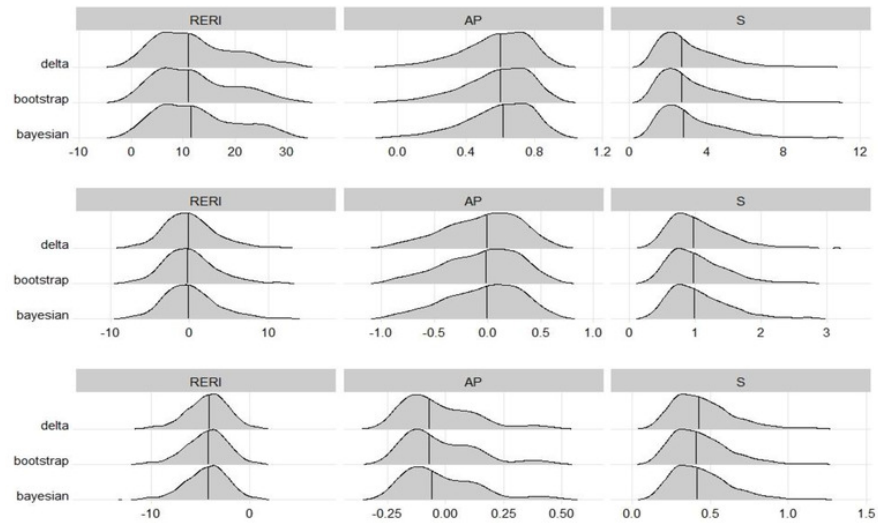



Figure 2

Figure 2. distributions of interaction in Oral cancer example.

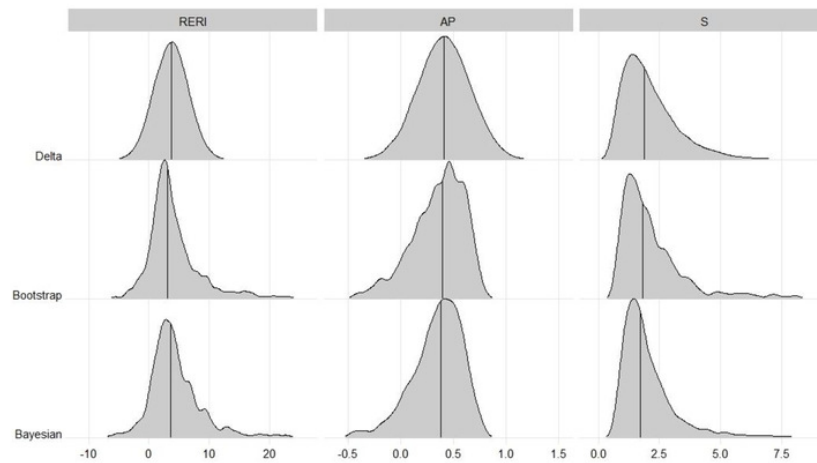
The distributions for Delta were based on the normal distribution (RERI and AP) or lognormal distribution (S) with estimated mean and standard error, respectively.

The ORs setting for simulation.

Table 1.   The ORs setting for simulation

scenario	OR_{01}	OR_{10}	OR_{11}	RERI	AP	S
A1	4.0	5.0	20.000	12.00	0.60	2.71
A2	4.0	5.0	12.000	4.00	0.33	1.57
A3	4.0	5.0	8.000	0.00	0.00	1.00
A4	4.0	5.0	6.000	-2.00	-0.33	0.71
A5	4.0	5.0	4.000	-4.00	-1.00	0.43
B1	2.0	5.0	15.000	9.00	0.60	2.80
B2	2.0	5.0	9.000	3.00	0.33	1.60
B3	2.0	5.0	6.000	0.00	0.00	1.00
B4	2.0	5.0	4.500	-1.50	-0.33	0.70
B5	2.0	5.0	3.000	-3.00	-1.00	0.40
C1	4.0	2.5	13.750	8.25	0.60	2.83
C2	4.0	2.5	8.250	2.75	0.33	1.61
C3	4.0	2.5	5.500	0.00	0.00	1.00
C4	4.0	2.5	4.125	-1.38	-0.33	0.69
C5	4.0	2.5	2.750	-2.75	-1.00	0.39
D1	2.0	2.5	8.750	5.25	0.60	3.10
D2	2.0	2.5	5.250	1.75	0.33	1.70
D3	2.0	2.5	3.500	0.00	0.00	1.00
D4	2.0	2.5	2.625	-0.88	-0.33	0.65
D5	2.0	2.5	1.750	-1.75	-1.00	0.30

Table 2. Coverage properties of the 95% two-sided confidence intervals for RERI in balanced design study.

1 Table 2. Coverage properties of the 95% two-sided confidence/credible intervals for RERI in balanced

2 design study

RERI	Delta				Bootstrap				Bayesian			
	Estimate	Left	Cover	Right	Estimate	Left	Cover	Right	Estimate	Left	Cover	Right
12.00	11.95	7.3	92.7	0.0	12.30	1.3	94.4	4.3	12.93	2.1	93.8	4.1
4.00	3.82	5.2	94.8	0.0	3.78	2.1	95.3	2.6	4.07	2.8	94.8	2.4
0.00	−0.17	3.2	96.8	0.0	−0.24	2.9	94.8	2.3	−0.13	3.1	94.9	2.0
−2.00	−2.11	1.7	98.2	0.1	−2.17	3.1	94.8	2.1	−2.15	3.6	94.6	1.8
−4.00	−4.11	0.7	97.9	1.4	−4.18	3.3	94.7	2.0	−4.22	3.4	94.5	2.1
9.00	9.03	8.8	91.2	0.0	9.21	2.9	92.5	4.6	9.63	3.3	92.8	3.9
3.00	2.81	6.4	93.6	0.0	2.83	2.8	94.4	2.8	2.94	3.5	94.3	2.2
0.00	−0.09	3.4	96.6	0.0	−0.10	3.4	94.1	2.5	−0.08	3.8	94.1	2.1
−1.50	−1.52	1.4	98.4	0.2	−1.55	3.4	94.7	1.9	−1.56	3.0	95.0	2.0
−3.00	−3.03	0.7	97.2	2.1	−3.05	3.7	94.0	2.3	−3.11	4.0	93.7	2.3
8.25	8.44	7.5	92.5	0.0	8.48	1.7	94.0	4.3	8.94	2.4	93.8	3.8
2.75	2.70	4.6	95.4	0.0	2.67	2.4	95.0	2.6	2.82	3.5	94.3	2.2
0.00	−0.06	1.7	98.3	0.0	−0.09	2.7	95.3	2.0	−0.07	3.6	94.3	2.1
−1.38	−1.35	0.5	99.1	0.4	−1.39	2.7	95.4	1.9	−1.38	3.2	94.7	2.1
−2.75	−2.80	0.4	96.9	2.7	−2.85	3.3	94.8	1.9	−2.88	3.5	94.1	2.4
5.25	5.13	7.5	92.5	0.0	5.17	2.7	94.4	2.9	5.37	3.6	93.9	2.5
1.75	1.70	6.2	93.8	0.0	1.66	3.6	93.5	2.9	1.70	4.6	92.9	2.5
0.00	−0.04	3.1	96.9	0.0	−0.06	3.1	95.2	1.7	−0.06	3.4	95.1	1.5
−0.88	−0.87	1.9	97.9	0.2	−0.89	3.8	94.6	1.6	−0.91	3.7	94.7	1.6
−1.75	−1.80	0.9	96.7	2.4	−1.83	4.4	93.8	1.8	−1.84	4.7	93.9	1.4

Table 3. Coverage properties of the 95% two-sided confidence intervals for AP in balanced design study.

1 Table 3. Coverage properties of the 95% two-sided confidence intervals for AP in balanced design

2 study

	Delta				Bootstrap				Bayesian			
AP	Estimate	Left	Cover	Right	Estimate	Left	Cover	Right	Estimate	Left	Cover	Right
0.60	0.60	0.0	92.2	7.8	0.60	1.5	94.5	4.0	0.61	2.7	93.7	3.6
0.33	0.33	0.2	93.2	6.6	0.32	2.2	95.3	2.5	0.34	2.7	95	2.3
0.00	−0.02	0.6	93.3	6.1	−0.03	2.9	94.8	2.3	−0.02	3.1	94.9	2.0
−0.33	−0.37	0.6	93.7	5.7	−0.38	3.1	94.8	2.1	−0.37	3.3	94.8	1.9
−1.00	−1.04	0.3	93.1	6.6	−1.08	3.0	94.7	2.3	−1.06	3.4	94.5	2.1
0.60	0.60	0.6	91.8	7.6	0.61	3.0	93.2	3.8	0.62	3.5	93.2	3.3
0.33	0.33	0.7	92.5	6.8	0.32	3.1	94.1	2.8	0.33	3.4	94.4	2.2
0.00	−0.01	0.6	93.5	5.9	−0.02	3.4	94.1	2.5	−0.01	3.8	94.1	2.1
−0.33	−0.34	0.4	92.5	7.1	−0.36	3.1	95.0	1.9	−0.35	3.1	94.8	2.1
−1.00	−1.03	0.1	92.4	7.5	−1.04	3.7	93.6	2.7	−1.05	3.4	94.0	2.6
0.60	0.60	0.1	91.5	8.4	0.61	1.8	94.6	3.6	0.62	2.8	93.9	3.3
0.33	0.34	0.1	93.8	6.1	0.33	2.4	95.0	2.6	0.34	3.0	94.4	2.6
0.00	−0.01	0.2	92.7	7.1	−0.02	2.8	95.2	2.0	−0.01	3.6	94.3	2.1
−0.33	−0.34	0.2	92.7	7.1	−0.35	2.9	95.6	1.5	−0.35	3.0	95.3	1.7
−1.00	−1.00	0.0	93.1	6.9	−1.04	3.2	95.0	1.8	−1.05	3.3	94.8	1.9
0.60	0.60	1.0	92.4	6.6	0.60	3.2	94.9	1.9	0.60	3.8	94.2	2.0
0.33	0.33	0.8	92.4	6.8	0.33	3.5	94.1	2.4	0.32	3.8	94.2	2.0
0.00	−0.01	1.1	91.3	7.6	−0.02	3.1	95.2	1.7	−0.02	3.4	95.1	1.5
−0.33	−0.33	0.8	91.3	7.9	−0.35	3.1	95.5	1.4	−0.34	3.3	95.1	1.6
−1.00	−1.01	1.0	90.3	8.7	−1.04	3.7	94.1	2.2	−1.02	3.2	94.8	2.0

Table 4. Coverage properties of the 95% two-sided confidence intervals for S in balanced design study.

2   Table 4. Coverage properties of the 95% two-sided confidence intervals for S in balanced design study

	Delta				Bootstrap				Bayesian			
S	Estimate	Left	Cover	Right	Estimate	Left	Cover	Right	Estimate	Left	Cover	Right
2.71	2.70	3.1	96.8	0.1	2.73	1.5	94.7	3.8	2.84	2.6	93.5	3.9
1.57	1.56	2.8	95.8	1.4	1.56	2.3	95.2	2.5	1.59	2.5	95.2	2.3
1.00	0.98	3.0	95.8	1.2	0.97	2.9	94.8	2.3	0.98	3.1	94.9	2.0
0.71	0.70	2.6	95.8	1.6	0.69	3.2	94.7	2.1	0.69	3.4	94.7	1.9
0.43	0.42	2.1	95.5	2.4	0.41	3.1	94.5	2.4	0.42	3.5	94.4	2.1
2.80	2.84	3.8	95.8	0.4	2.89	3.0	93.1	3.9	2.93	3.7	93.0	3.3
1.60	1.57	3.6	95.1	1.3	1.58	3.2	93.9	2.9	1.59	3.4	94.3	2.3
1.00	0.98	3.3	94.8	1.9	0.97	3.4	94.1	2.5	0.99	3.8	94.1	2.1
0.70	0.69	2.1	96.1	1.8	0.68	2.9	95.2	1.9	0.69	3.2	94.7	2.1
0.40	0.39	1.0	96.2	2.8	0.38	3.4	94.1	2.5	0.39	2.9	94.5	2.6
2.83	2.90	3.3	95.6	1.1	2.97	2.0	94.8	3.2	3.04	2.8	94.1	3.1
1.61	1.61	2.8	95.9	1.3	1.62	2.3	95	2.7	1.63	2.9	94.6	2.5
1.00	0.98	2.5	95.8	1.7	0.98	2.8	95.2	2.0	0.99	3.6	94.3	2.1
0.69	0.69	1.9	96.6	1.5	0.68	2.8	95.6	1.6	0.69	3.0	95.3	1.7
0.39	0.39	0.7	97.2	2.1	0.38	2.8	95.6	1.6	0.38	3.3	95.0	1.7
3.10	3.11	4.6	94.6	0.8	3.15	2.9	95.6	1.5	3.15	3.9	94.5	1.6
1.70	1.68	3.5	95.2	1.3	1.70	3.5	94.6	1.9	1.68	3.7	94.5	1.8
1.00	0.98	2.4	96.4	1.2	0.97	3.1	95.4	1.5	0.98	3.4	95.1	1.5
0.65	0.65	0.7	97.8	1.5	0.64	2.9	95.4	1.7	0.64	2.9	95.6	1.5
0.30	0.29	0.0	96.7	3.3	0.28	2.7	95.2	2.1	0.29	1.9	95.5	2.6

Table 5. Distribution of exposures among cases and controls in the Oral cancer example[a]

[a] data were from the example presented by Rothman and Keller

1   Table 5. Distribution of exposures among cases and controls in the Oral cancer example[a]

	Neither	Smoking only	Alcohol only	Smoking and Alcohol
Cases	3	8	6	225
Controls	20	18	12	166

2   [a] data were from the example presented by Rothman and Keller

Table 6. Point estimates with 95% confidence limits for interaction effects in Oral cancer example

2   Table 6. Point estimates with 95% confidence limits for interaction effects in Oral cancer example

	Delta	Bootstrap	Bayesian
RERI	3.74 (−1.84 to 9.32)	3.51 (−4.86 to 5.78×10^{6})	3.68 (−3.93 to 19.59)
AP	0.41 (−0.07 to 0.90)	0.38 (−0.32 to 0.83)	0.38 (−0.34 to 0.72)
S	1.87 (0.65 to 5.42)	1.69 (0.66 to 11.24)	1.71 (0.72 to 5.78)