Published in final edited form as:
Epidemiol Method. 2012 August 1; 1(1): 159–188. doi:10.1515/2161-962X.1010.

Sample Size and Power Calculations for Additive Interactions

T.J. VanderWeele
Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston, MA
T.J. VanderWeele: tvanderw@hsph.harvard.edu

Introduction

The literature on power and sample size calculations for interaction has focused on the multiplicative scale (Lubin and Gail, 1990; Hwang et al., 1994; Foppa and Spiegelman, 1997; Yang et al., 1999; Garcia-Closas and Lubin, 1999; Qiu et al., 2000; Luan et al. 2001; Gauderman, 2002a b; Sturmer, 2002; Wang et al., 2003; Wang and Zhao, 2003; Demidenko, 2008; VanderWeele, 2011). However, interaction on the additive scale is more relevant for public health purposes (Rothman et al.,1980; Rothman et al., 2008) and is also more closely related to notions of mechanistic synergism within the sufficient cause framework (Rothman, 1976; VanderWeele and Robins, 2007, 2008; Rothman et al., 2008). Arguably, the reason interaction is most frequently assessed on the multiplicative scale is that this is what is most easily computed from the output of standard logistic regression software. In addition, in the context of case-control studies, odds ratios can be estimated but risk differences cannot, unless additional information concerning e.g. the prevalence of the outcome or exposures in the underlying population is available (Rothman et al., 2008). This again leads to the multiplicative scale as being the default for assessing interaction. That power and sample size calculations are better developed for multiplicative interaction than for additive interaction perhaps further encourages the use of the multiplicative scale for interaction assessment. However, measures of additive interaction based on risk ratios or odds ratios using the relative excess risk due to interaction ("RERI"; Rothman, 1986) can easily be calculated from logistic regression with either cohort or case-control data (Hosmer and Lemeshow, 1992) and in this paper we will derive power and sample size formulae for interaction on the additive scale. Power and sample size calculations for additive interaction were discussed in Greenland (1983, 1985) but no closed form expressions were provided.

In this paper, we will consider measures of additive interaction based on absolute risks and also on the relative excess risk due to interaction for both cohort and case-control data and we will provide closed form analytic expressions for power and sample size in each of these cases. Analytically, we will for the most part follow the development of Demidenko (2008) who considered multiplicative interaction but we will be taking a similar approach for the additive scale. We will see that when main effects of both exposures are positive, power to detect positive interaction on the additive scale will be greater than that on the multiplicative scale, providing yet another reason, beyond public health relevance and relation to mechanistic synergism, for using the additive scale to assess interaction. The reader who is

interested in only the application of the power and sample size formulae derived in this paper is referred to Appendix 2 at the end of the paper on epidemiologic practice. This appendix gives instructions on using Excel spreadsheets (included as an online supplement to this paper) to automatically carry out power and sample size calculations for additive and multiplicative interaction for cohort, case-control and case-only data.

Notation and Definitions

We will suppose we have a binary outcome Y and two binary exposures G and E. Although G and E might represent genetic and environmental exposures, respectively, nothing in the development will require this. They might be two environmental exposures, or two genetic exposures, or behavior exposures, etc. Let $p_{ge} = P(Y = 1|G = g, E = e)$ and let $\pi_{ge} = P(G = g, E = e)$. The measure of interaction on the additive scale using risks is then

$$p_{11} - p_{10} - p_{01} + p_{00}.$$

This can be re-expressed as $(p_{11} - p_{00}) - \{(p_{10} - p_{00}) + (p_{01} - p_{00})\}$ and measures the extent to which the effect of both exposures combined exceeds (or is less than) the sum of the effects of each exposure considered separately. If $p_{11} - p_{10} - p_{01} + p_{00} > 0$, the interaction is said to be positive or "superadditive". If $p_{11} - p_{10} - p_{01} + p_{00} < 0$, the interaction is said to be negative or "subadditive". If $p_{11} - p_{10} - p_{01} + p_{00} = 0$, there is said to be no interaction on the additive scale. This measure of additive interaction corresponds to the coefficient of the product term for the two exposure in a linear risk model for the outcome.

In many studies, analyses are presented using risk ratios or odds ratios rather than absolute risks. Define the risk ratio as $RR_{ge} = \frac{P(Y=1|G=g,E=e)}{P(Y=1|G=0,E=0)} = \frac{p_{ge}}{p_{00}}$ and the odds ratio as $OR_{ge} = \frac{P(Y=1|G=g,E=e)/P(Y=0|G=g,E=e)}{P(Y=1|G=0,E=0)/P(Y=0|G=0,E=0)} = \frac{p_{ge}/(1-p_{ge})}{p_{00}/(1-p_{00})}$. The measure of multiplicative interaction used on the risk ratio or odds ratio scale is then generally taken as $I_{RR} = \frac{RR_{11}}{RR_{10}RR_{01}}$ or $I_{OR} = \frac{OR_{11}}{OR_{10}OR_{01}}$ respectively. These measures of multiplicative interaction correspond to the exponentiated coefficients of the product term for the two exposures in log-linear and logistic regression models for the outcome respectively.

Suppose now we were to divide our measure of additive interaction based on risks, $p_{11} - p_{10} - p_{01} + p_{00}$, by the baseline risk p_{00}. We would then obtain what is sometimes referred to as the relative excess risk due to interaction or *RERI* (Rothman, 1986):

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

This measure *RERI* will be greater than 0 (or respectively less than 0) if and only if the measure of additive interaction using absolute risks, $p_{11} - p_{10} - p_{01} + p_{00}$, is greater than 0 (or less than 0 respectively). The relative excess risk due to interaction can thus be used to assess additive interaction using data on relative risks. When the probability of the outcome

is rare in all exposure strata then odds ratios will approximate risk ratios i.e.

$\frac{p_{ge}/(1-p_{ge})}{p_{00}/(1-p_{00})} \approx \frac{p_{ge}}{p_{00}}$ and thus we can approximate *RERI* by

$$RERI_{OR}=OR_{11}-OR_{10}-OR_{01}+1 \approx RERI.$$

This final measure, $RERI_{OR} = OR_{11} - OR_{10} - OR_{01} + 1$, is advantageous because it is an approximate measure of additive interaction and yet can also be obtained directly from logistic regression analyses and from case-control data. We will, however, first begin with additive interaction on the absolute risk scale using cohort data.

Additive Interaction in Cohort Studies Using A Linear Risk Model

Suppose data were available from a cohort study and we were to use a linear risk model to measure additive interaction:

$$P(Y=1|G=g,E=e)=\theta_0+\theta_1 g+\theta_2 e+\theta_3 ge. \quad (1)$$

In this model $\theta_3 = p_{11} - p_{10} - p_{01} + p_{00}$ is our measure of additive interaction. Suppose we plan to fit this model to the cohort data using maximum likelihood and use a Wald test for the null hypothesis $\theta_3 = 0$. Once we have fit the model and obtained an estimate $\hat{\theta_3}$ of θ_3 from the data, the Wald test statistic for the null hypothesis $\theta_3 = 0$ is given by $\hat{\theta_3}/\hat{V}$ where $\hat{V}$ is the estimated variance of $\hat{\theta_3}$. We would reject the null at significance level α if $|\hat{\theta_3}/\hat{V}| > Z_{1-\alpha/2}$ where $Z_{1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard normal distribution. Suppose we wish to calculate the sample size required to reject the null hypothesis with significance level α and power β if the magnitude of the interaction were $\theta_3 = \eta$.

By standard sample size arguments, the sample size required to detect an additive interaction of magnitude $\theta_3 = \eta$ with significance level α and power β is

$$n=\frac{(Z_{1-\alpha/2}+Z_\beta)^2 V}{\eta^2}$$

where $Z_{1-\alpha/2}$ and Z_β are the $(1-\alpha/2)$th and βth quantiles respectively of the standard normal distribution and where V is the variance of $\hat{\theta_3}$ under the alternative that $\theta_3 = \eta$. The difficulty lies in calculating the variance V. In Appendix 1, we show that the variance V is given by

$$V=\frac{1}{L'}+\frac{1}{F'}+\frac{1}{J'}+\frac{1}{R'}$$

where

$$L' = \frac{1}{(\theta_0)(1-\theta_0)}\pi_{00}$$
$$F' = \frac{1}{(\theta_0+\theta_1)\{1-(\theta_0+\theta_1)\}}\pi_{10}$$
$$J' = \frac{1}{(\theta_0+\theta_2)\{1-(\theta_0+\theta_2)\}}\pi_{01}$$
$$R' = \frac{1}{(\theta_0+\theta_1+\theta_2+\theta_3)\{1-(\theta_0+\theta_1+\theta_2+\theta_3)\}}\pi_{11}.$$

Thus to calculate the sample size we would need to specify (i) the significance level α, the power β, and the magnitude of additive interaction $\theta_3 = \eta$; (ii) the proportion of subjects in each exposure stratum, π_{00}, π_{10}, π_{01}, π_{11}; and (iii) the main effect of the two exposures on the additive scale θ_1 and θ_2 and the baseline risk of the doubly unexposed group $\theta_0 = P(Y = 1|G = 0, E = 0)$.

Instead of specifying the proportion of subjects in each joint exposure stratum π_{00}, π_{10}, π_{01}, π_{11}, we could instead specify the marginal probability of each exposure $\pi_g = P(G = 1)$ and $\pi_e = P(E = 1)$ along with the odds ratio relating G and E, $\Delta = \{P(G = 1|E = 1)/P(G = 0|E = 1\}/\{P(G = 1|E = 0)/P(G = 0|E = 0\}$. The probabilities π_{00}, π_{10}, π_{01}, π_{11} are then given by (Demidenko, 2008):

$$\begin{aligned}\pi_{00} &= \frac{1-\pi_e}{1+C}\\ \pi_{10} &= \frac{(1-\pi_e)C}{1+C}\\ \pi_{01} &= \frac{\pi_e}{1+C\Delta}\\ \pi_{11} &= \frac{C\Delta\pi_e}{1+C\Delta}\end{aligned} \quad (2)$$

where

$$C = \frac{q+\sqrt{q^2+4\pi_g(1-\pi_g)\Delta}}{2(1-\pi_g)\Delta}$$
$$\text{and where } q = \pi_g(1+\Delta)+\pi_e(1-\Delta)-1.$$

If G and E are independent then $\Delta = 1$ and C simplifies to $C = \pi_e/(1 - \pi_e)$.

If instead of calculating the required sample size for a fixed power β, we wanted to calculate the power for a given sample size using the Wald test for the null hypothesis $\theta_3 = 0$ based on model (1) we could proceed as follows. For a fixed sample size n the power to reject the null $\theta_3 = 0$ at significance level α under the alternative that $\theta_3 = \eta$ is given by

$$Power = \Phi^{-1}\left\{-Z_{1-\alpha/2}+\eta\sqrt{(n/V)}\right\}$$

where Φ^{-1} is the inverse cumulative distribution function for a standard normal random variable and where V can be calculated as above. In Appendix 2 we describe how to use a simple Excel spreadsheet (included with this paper as an online supplement) to carry out such sample size and power calculations automatically. The online supplement also provides Excel spreadsheets for the sample size and power calculations for additive interaction using

relative excess risk due to interaction from logistic regression with cohort or case-control data described in the following sections. The use of these Excel spreadsheets is described in detail in Appendix 2. Finally, it should be noted that if the null hypothesis were rejected for extreme values of θ_3 on either side of zero (two-sided test) then the relevant power formula would be:

$$Power=\Phi^{-1}\left\{-Z_{1-\alpha/2}+\eta\sqrt{(n/V)}\right\}+\Phi^{-1}\left\{-Z_{1-\alpha/2}-\eta\sqrt{(n/V)}\right\}.$$

Before moving on, we give a brief example of the use of these formulae for additive interaction.

Example 1

Suppose we wish to calculate the power of a test at significance level $\alpha = 0.05$, with $n = 4000$, with the prevalence of the genetic and environmental factors being $\pi_g = 0.5$ and $\pi_e = 0.3$ respectively and assuming these are independent so that $\Delta = 1$, with the probability of the outcome in the reference category of $\theta_0 = P(Y = 1|G = 0, E = 0) = 0.02$, with main effects on the risk difference scale of $\theta_1 = 0.01$ and $\theta_2 = 0.01$ and with additive interaction $\theta_3 = 0.02$. We can use the equations in (2) to calculate $\pi_{00} = 0.35$, $\pi_{10} = 0.35$, $\pi_{01} = 0.15$, $\pi_{11} = 0.15$ and from this we can calculate L', F', J', R' and the variance V and the power $Power=\Phi^{-1}\left\{-Z_{1-\alpha/2}+\eta\sqrt{(n/V)}\right\}$ to obtain 0.32.

Additive Interaction in Cohort Studies Using Logistic Regression and RERI

In this section we consider power and sample size calculations for measures of interaction based on $RERI_{OR}$ obtained from logistic regression using cohort data. We will first review the power and sample size calculations for multiplicative interaction from logistic regression using cohort data given by Demidenko (2008) since the variance calculation of Demidenko will underlie those given here for additive interaction using the relative excess risk due to interaction.

Suppose we fit a logistic regression model to cohort data:

$$\text{logit}\{P(Y=1|G=g,E=e)\}=\gamma_0+\gamma_1 g+\gamma_2 e+\gamma_3 ge. \quad (3)$$

The coefficient γ_3 is generally referred to as a measure of interaction of the multiplicative scale. The exponentiated coefficient is equal to the odds ratio multiplicative interaction ratio $e^{\gamma_3}=I_{OR}=\frac{OR_{11}}{OR_{10}OR_{01}}$. Suppose we wish to use a Wald test for the null hypothesis $\gamma_3 = 0$. The sample size required to detect a multiplicative interaction of magnitude $\gamma_3 = \eta$ with significance level α and power β is

$$n=\frac{(Z_{1-\alpha/2}+Z_{\beta})^2 V_{mult(OR)}}{\eta^2}$$

where $Z_{1-\alpha/2}$ and Z_{β} are the $(1-\alpha/2)$th and βth quantiles respectively of the standard normal distribution and where $V_{mult(OR)}$ is the variance of $\hat{\gamma_3}$ under the alternative that $\theta_3 = \eta$. Demidenko (2008) derives the variance matrix for the maximum likelihood estimator of $(\gamma_0, \gamma_1, \gamma_2, \gamma_3)$, given in Appendix 1, and specifically shows that

$$V_{mult(OR)}=\frac{1}{L}+\frac{1}{F}+\frac{1}{J}+\frac{1}{R}$$

where

$$\begin{aligned} L&=\frac{e^{\gamma_0}}{(1+e^{\gamma_0})^2}\pi_{00} \\ F&=\frac{e^{\gamma_0+\gamma_1}}{(1+e^{\gamma_0+\gamma_1})^2}\pi_{10} \\ J&=\frac{e^{\gamma_0+\gamma_2}}{(1+e^{\gamma_0+\gamma_2})^2}\pi_{01} \\ R&=\frac{e^{\gamma_0+\gamma_1+\gamma_2+\gamma_3}}{(1+e^{\gamma_0+\gamma_1+\gamma_2+\gamma_3})^2}\pi_{11}. \end{aligned} \quad (4)$$

Once again, to calculate the sample size we would need to specify (i) the significance level α, the power β, and the magnitude of additive interaction $\gamma_3 = \eta$; (ii) the proportion of subjects in each exposure stratum, π_{00}, π_{10}, π_{01}, π_{11}; and (iii) the main effect odds ratios of the two exposures on the logistic scale, γ_1 and γ_2, and the log odds of the baseline risk of the doubly unexposed group $\gamma_0 = \log\{P(Y = 1|G = 0, E = 0)/P(Y = 0|G = 0, E = 0)\}$. Once again, if instead of specifying the joint probabilities π_{00}, π_{10}, π_{01}, π_{11}, we specified the marginal probabilities of each exposure $\pi_g = P(G = 1)$ and $\pi_e = P(E = 1)$ and the odds ratio relating G and E, $\Delta = \{P(G = 1|E = 1)/P(G = 0|E = 1\}/\{P(G = 1|E = 0)/P(G = 0|E = 0\}$ then we could obtain the π_{00}, π_{10}, π_{01}, π_{11} using the formulae in (2). And once again, if instead of calculating the required sample size for a given power, we wanted to calculate the power for a given sample size we could use $Power=\Phi^{-1}\left\{-Z_{1-\alpha/2}+\eta\sqrt{(n/V_{mult(OR)})}\right\}$.[1]

[1]Demidenko (2008) also noted that a number of previous authors (Hwang et al., 1994; Foppa and Spiegelman, 1997) who had considered sample size and power calculations for interaction in logistic regression had relied on a different formula for their sample size calculations. These other authors had assumed that for the test statistic, the variance of $\hat{\gamma_3}$ had been calculated under the null hypothesis of no interaction. When the variance for the test statistic is calculated under the null of no interaction then the required sample size is given by $n=\frac{(Z_{1-\alpha/2}\sqrt{V_0}+Z_{\beta}\sqrt{V_{mult(OR)}})}{\eta^2}$ rather than by $n=\frac{(Z_{1-\alpha/2}+Z_{\beta})^2 V_{mult(OR)}}{\eta^2}$ where V_0 is the variance of $\hat{\gamma_3}$ calculated under the null that $\gamma_3 = 0$. Demidenko (2008) points out that although the sample size calculations of Hwang et al. (1994) and Foppa and Spiegelman (1997) would be fine if, for $\hat{\gamma_3}$, the variance were indeed calculated under the null, in practice, the variance of $\hat{\gamma_3}$ is almost always calculated under the alternative; it is the variance under the alternative that is generally given as the default in standard logistic regression output. Thus, the sample size calculations of Hwang et al. (1994) and Foppa and Spiegelman (1997), although not technically incorrect, do not correspond to the test statistics that are generally used in practice. A similar point and criticism was made by Garcia-Closas and Lubin (1999) some years earlier. Both Garcia-Closas and Lubin (1999) and Demidenko (2008) note that when interactions are large, the sample size calculations using the "null-variance" can underestimate the required sample size if the test statistic with the variance under the alternative is in fact used. Likewise a similar point pertains to the sample size and power calculations of Yang et al. (1997) for multiplicative interaction in case-only studies (cf. VanderWeele, 2011).

Example 2

Suppose we wish to calculate the power of a test at significance level $\alpha = 0.05$, with $n = 5000$, with the joint prevalence of the genetic and environmental factors being $\pi_{00} = 0.35$, $\pi_{10} = 0.20$, $\pi_{01} = 0.20$, $\pi_{11} = 0.25$ respectively, with the probability of the outcome in the reference category of $P(Y = 1|G = 0, E = 0) = 0.015$, with main effects on the odds ratio scale of $e^{\gamma 1} = 1.3$ and $e^{\gamma 2} = 1.4$ and with odds ratio multiplicative interaction $e^{\gamma 3} = 1.6$. We can calculate L, F, J, R from these values and the variance $V_{mult(OR)}$ to obtain

$$Power = \Phi^{-1}\left\{-Z_{1-\alpha/2} + \eta\sqrt{(n/V)}\right\} = 0.216.$$

We will now use the variance matrix calculations of Demidenko (2008) to derive sample size and power formulae for the relative excess risk due to interaction (RERI). The RERI from logistic regression model (3) is given by:

$$RERI_{OR} = e^{\gamma_1+\gamma_2+\gamma_3} - e^{\gamma_1} - e^{\gamma_2} + 1.$$

Suppose we wish to use a Wald test for the null hypothesis $RERI_{OR} = 0$. The sample size required to detect a $RERI_{OR}$ of magnitude $\eta = e^{\gamma 1+\gamma 2+\gamma 3} - e^{\gamma 1} - e^{\gamma 2} + 1$ with significance level α and power β is

$$n = \frac{(Z_{1-\alpha/2} + Z_\beta)^2 V_{RERI(OR)}}{\eta^2}$$

where $Z_{1-\alpha/2}$ and Z_β are the $(1-\alpha/2)$th and βth quantiles respectively of the standard normal distribution and where $V_{RERI(OR)}$ is the variance of $\widehat{RERI}_{OR} = e^{\hat{\gamma}1+\hat{\gamma}2+\hat{\gamma}3} - e^{\hat{\gamma}1} - e^{\hat{\gamma}2} + 1$ under the alternative. Using the delta method, we show in Appendix 1 that this variance is given by:

$$V_{RERI(OR)} = \left(\frac{1}{L}+\frac{1}{R}\right)e^{2(\gamma_1+\gamma_2+\gamma_3)} - \frac{2}{L}e^{2\gamma_1+\gamma_2+\gamma_3} - \frac{2}{L}e^{\gamma_1+2\gamma_2+\gamma_3} + \left(\frac{1}{L}+\frac{1}{F}\right)e^{2\gamma_1} + \left(\frac{1}{L}+\frac{1}{J}\right)e^{2\gamma_2} + \frac{2}{L}e^{\gamma_1+\gamma_2}$$

where L, F, J, R are given as in equation (4) above.

To calculate the sample size to reject the null of no additive interaction using $RERI_{OR}$, we would need to specify (i) the significance level α, the power β; (ii) the proportion of subjects in each exposure stratum, π_{00}, π_{10}, π_{01}, π_{11}; and (iii) the main effect odds ratios of the two exposures on the logistic scale, γ_1 and γ_2, the log odds of the baseline risk of the doubly unexposed group $\gamma_0 = \log\{P(Y = 1|G = 0, E = 0)/P(Y = 0|G = 0, E = 0)\}$, and the magnitude of the interaction on the multiplicative scale γ_3. Instead of specifying the magnitude of the interaction on the multiplicative scale, γ_3, one could specify the magnitude of $RERI_{OR}$ under the alternative $RERI_{OR} = \eta$ and then back-calculate the magnitude of $\gamma_3 = \log(\eta + e^{\gamma 1} - e^{\gamma 2} - 1) - \gamma_1 - \gamma_2$.

And once again, if instead of specifying the joint probabilities π_{00}, π_{10}, π_{01}, π_{11}, we specified the marginal probabilities of each exposure $\pi_g = P(G = 1)$ and $\pi_e = P(E = 1)$ and the odds ratio relating G and E, $\Delta = \{P(G = 1|E = 1)/P(G = 0|E = 1\}/\{P(G = 1|E = 0)/P(G = 0|E = 0\}$ then we could obtain π_{00}, π_{10}, π_{01}, π_{11} using the formulae in (2). And once again, if instead of calculating the required sample size for a given power, we wanted to calculate the power for a given sample size we could use $Power = \Phi^{-1}\left\{-Z_{1-\alpha/2} + \eta\sqrt{n/V_{RERI(OR)}}\right\}$.

Example 3

Suppose again we wish to calculate the power of a test at significance level $\alpha = 0.05$, with $n = 5000$, with the joint prevalence of the genetic and environmental factors being $\pi_{00} = 0.35$, $\pi_{10} = 0.20$, $\pi_{01} = 0.20$, $\pi_{11} = 0.25$ respectively, with the probability of the outcome in the reference category of $P(Y = 1|G = 0, E = 0) = 0.015$, with main effects on the odds ratio scale of $e^{\gamma 1} = 1.3$ and $e^{\gamma 2} = 1.4$ and with odds ratio multiplicative interaction $e^{\gamma 3} = 1.6$ as in Example 2, but that we now wish to calculate the power for testing $RERI_{OR} > 0$. Here the true $RERI_{OR}$ is $\eta = e^{\gamma 1+\gamma 2+\gamma 3} - e^{\gamma 1} - e^{\gamma 2} + 1 = (1.3)(1.4)(1.6) - (1.3) - (1.4) + 1 = 1.212 > 0$. From L, F, J, R we can calculate the variance $V_{RERI(OR)}$ to obtain $Power = \Phi^{-1}\left\{-Z_{1-\alpha/2} + \eta\sqrt{(n/V_{RERI(OR)})}\right\} = 0.482$. In this example, the power to detect additive interaction, 0.482, is greater than that to detect multiplicative interaction, 0.216.

The reader is reminded that the tests for additive interaction using $RERI_{OR}$ hold only approximately to the extent that the outcome is rare so that $RERI_{OR}$ approximates *RERI* on the risk ratio scale. In Appendix 1 we also derive sample size and power formulae for the multiplicative interaction from a log-linear model and for additive interaction using *RERI* estimated from a log-linear model. However, if the measure of additive interaction is fit with cohort data, it may be preferable to fit model (1) directly for additive interaction using absolute risks rather than employing *RERI*.

Additive Interaction in Case-Control Studies Using Logistic Regression and RERI

Suppose instead we fit a logistic regression model to case-control data:

$$\log it\{P(Y = 1|G = g, E = e)\} = \gamma_0 + \gamma_1 g + \gamma_2 g + \gamma_3 ge.$$

The sample size required to detect a $RERI_{OR}$ of magnitude $\eta = e^{\gamma 1+\gamma 2+\gamma 3} - e^{\gamma 1} - e^{\gamma 2} + 1$ with significance level α and power β is

$$n = \frac{(Z_{1-\alpha/2} + Z_\beta)^2 V^*_{RERI(OR)}}{\eta^2}$$

where

$$V^*_{RERI(OR)}=(\frac{1}{L^*}+\frac{1}{R^*})e^{2(\gamma_1+\gamma_2+\gamma_3)}-\frac{2}{L^*}e^{2\gamma_1+\gamma_2+\gamma_3}-\frac{2}{L^*}e^{\gamma_1+2\gamma_2+\gamma_3}+(\frac{1}{L^*}+\frac{1}{F^*})e^{2\gamma_1}+(\frac{1}{L^*}+\frac{1}{J^*})e^{2\gamma_2}+\frac{2}{L^*}e^{\gamma_1+\gamma_2}$$

with

$$\begin{aligned}L^*&=\frac{e^{\gamma_0}}{(1+e^{\gamma_0})^2}\pi^*_{00}\\ F^*&=\frac{e^{\gamma_0+\gamma_1}}{(1+e^{\gamma_0+\gamma_1})^2}\pi^*_{10}\\ J^*&=\frac{e^{\gamma_0+\gamma_2}}{(1+e^{\gamma_0+\gamma_2})^2}\pi^*_{01}\\ R^*&=\frac{e^{\gamma_0+\gamma_1+\gamma_2+\gamma_3}}{(1+e^{\gamma_0+\gamma_1+\gamma_2+\gamma_3})^2}\pi^*_{11}.\end{aligned}$$

and where $\pi^*_{00}, \pi^*_{10}, \pi^*_{01}, \pi^*_{11}$ are now the proportions of subjects in each joint exposure stratum *in the case-control sample*.

If we know the overall outcome prevalence in the underlying population, $P(Y = 1)$, we could also obtain the proportions $\pi^*_{00}, \pi^*_{10}, \pi^*_{01}, \pi^*_{11}$ from the proportions of subjects in each joint exposure stratum in the underlying population, π_{00}, π_{10}, π_{01}, π_{11}, though doing so requires solving a non-linear equation numerically (Demidenko, 2008). Alternatively, if the outcome is rare we can obtain $\pi^*_{00}, \pi^*_{10}, \pi^*_{01}, \pi^*_{11}$ from π_{00}, π_{10}, π_{01}, π_{11} approximately using the following formulas (see Appendix 1 for proof):

$$\begin{aligned}\pi^*_{00} &\approx \pi_{00}P^*(Y=0)+\frac{\pi_{00}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)\\ \pi^*_{10} &\approx \pi_{10}P^*(Y=0)+\frac{e^{\gamma_1}\pi_{10}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)\\ \pi^*_{01} &\approx \pi_{01}P^*(Y=0)+\frac{e^{\gamma_2}\pi_{01}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)\\ \pi^*_{11} &\approx \pi_{11}P^*(Y=0)+\frac{e^{\gamma_1+\gamma_2+\gamma_3}\pi_{11}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)\end{aligned}$$

where $P^*(Y = 0)$ is the proportion of controls in the case-control sample and $P^*(Y = 1)$ is the proportion of cases in the case-control sample. If we instead specify the marginal probabilities of each exposure $\pi_g = P(G = 1)$ and $\pi_e = P(E = 1)$ and the odds ratio, Δ, relating G and E, in the underlying population then we can calculate π_{00}, π_{10}, π_{01}, π_{11} using the formulae in (2).

Thus, to calculate the sample size to reject the null of no additive interaction using $RERI_{OR}$ from case-control data we would need to specify (i) the significance level α, the power β; (ii) the proportion of subjects in each exposure stratum, $\pi^*_{00}, \pi^*_{10}, \pi^*_{01}, \pi^*_{11}$ in the case-control sample, or alternatively these proportions π_{00}, π_{10}, π_{01}, π_{11} or the marginal probabilities and marginal odds ratio, π_g, π_e, Δ, n the underlying population along with a rare outcome assumption and the proportions of cases $P^*(Y = 1)$ in the case-control sample, and finally (iii) the main effect odds ratios of the two exposures on the logistic scale, γ_1 and γ_2, the log odds of the baseline probability of the outcome in the doubly unexposed group $\gamma_0 = \log\{P^*(Y = 1|G = 0, E = 0)/P^*(Y = 1|G = 0, E = 0)\}$ in the case-control sample, and the magnitude of the interaction on the multiplicative scale γ_3 (or instead the magnitude of $RERI_{OR} = \eta$ and then back-calculate the magnitude of $\gamma_3 = \log(\eta + e^{\gamma_1} - e^{\gamma_2} - 1) - \gamma_1 - \gamma_2$.

Note that if the joint or marginal exposure probabilities are specified separately for the cases and controls then under an assumption of a rare outcome, the distribution of the exposures amongst the controls could be used as an approximation to π_{00}, π_{10}, π_{01}, π_{11} or π_g, π_e, Δ.

Note also that with case control data, $\gamma_0 = \log\{P^*(Y = 1|G = 0, E = 0)/P^*(Y = 0|G = 0, E = 0)\}$ is the log odds of baseline probability of the outcome in doubly unexposed group *in the case-control sample* i.e. the log the number of cases to controls in the study for the doubly unexposed group. It is shown in the Appendix that under a rare outcome assumption γ_0 can be approximated as $\gamma_0 \approx \log it[1/\{1+(\pi_{00}+\pi_{10}e^{\gamma 1} + \pi_{01}e^{\gamma 2} + \pi_{11}e^{\gamma 1+\gamma 2+\gamma 3})P^*(Y = 0)/P^*(Y = 1)\}]$.

Example 4

Suppose we wish to calculate the same size required for a test at significance level $\alpha = 0.05$, with power $\beta = 0.80$, with the joint prevalence of the genetic and environmental factors being $\pi_g = 0.5$, $\pi_e = 0.3$ respectively in the underlying population with the factors being independent in the underlying population so that $\Delta = 1$. Suppose that the number of cases and controls in the study were going to be equal $P^*(Y = 1) = P^*(Y = 0) = 0.5$, with main effects on the odds ratio scale of $e^{\gamma 1} = 1.1$ and $e^{\gamma 2} = 1.1$ and with multiplicative interaction $e^{\gamma 3} = 1.5$. We can calculate that the sample size then required to detect positive multiplicative interaction would be $n = 3447$. We can also calculate that sample size required to detect positive interaction using $RERI_{OR}$ would be $n = 2212$.

It should also be noted that when multiplicative interaction is of interest and the genetic and environmental factors are independent of one another in the underlying population, a "case-only" estimator of multiplicative interaction will have greater power to detect multiplicative interaction as it exploits the independence assumption (Piegorsch et al., 1994; Yang et al., 1999). Power and sample size calculations for case-only estimators have been considered elsewhere (Yang et al., 1999; VanderWeele, 2011). Although these case-only estimators can be quite powerful, they are also fairly sensitive to the assumption that the two exposures are independent in the population and can result in considerable bias if this assumption does not hold (Albert et al., 2001).

A Power Comparison of Additive and Multiplicative Interaction

VanderWeele (2009a) noted that in a log-linear model with non-negative main effects, whenever positive multiplicative interaction is present on the risk ratio scale, positive additive interaction on the risk difference scale will be present as well; the reverse implication does not hold. Here we will explore power to detect such additive or multiplicative interaction and we will consider the odds ratio scale rather than the risk ratio scale. In this power comparison we will assume a case-control study with a rare outcome so that $RERI_{OR}$ approximates a measure of additive interaction. Table 1 below reports power for a number of scenarios with varying sample sizes, main effect odds ratios and multiplicative interaction parameters on the odds ratio scale $I_{OR} = \frac{OR_{11}}{OR_{10}OR_{01}}$.

In these examples it is assumed that the proportion of case and controls in the case-control sample are equal and that the prevalence of the genetic and environmental factors are each $\pi_g = \pi_e = 0.5$ with the odds ratio relating these factors being $\Delta = 1.1$. Note that in all scenarios considered there is positive interaction on both additive and multiplicative scales. Power for one-sided test (rejecting only for positive interaction) is reported.

We see that for the scenarios considered here with non-negative main effects and positive interaction, power is greater to detect additive interaction than multiplicative interaction. However, as noted in Greenland (1983), when outcome probabilities are additive or sub-additive, power to detect a (negative) multiplicative interaction will often be greater.

Power and Sample Size Calculations for Sufficient Cause Interactions and Epistatic Interactions

VanderWeele and Robins (2007, 2008) discuss “causal” or “sufficient cause” interactions within the sufficient cause and counterfactual frameworks (Rothman, 1976; Rubin, 1990; Hernán, 2004) which provide a somewhat stronger notion of positive additive interaction. A sufficient cause interaction is present if there are individuals for whom the outcome would occur if both exposures are present but would not occur if just one or the other exposure is present. In counterfactual notation, if we let Y_{ge} denote the counterfactual outcome (or potential outcome) for each subject if, possibly contrary to fact, G had been set to g and E had been set to e, then a sufficient cause interaction is present if for some individual $Y_{11} = 1$ but $Y_{10} = Y_{01} = 0$. VanderWeele and Robins (2007, 2008) showed that if the effect of the two exposures were un-confounded (in that the counterfactual outcomes Y_{ge} were independent of the actual exposures $\{G, E\}$) then

$$p_{11} - p_{10} - p_{01} > 0$$

would imply the presence of a sufficient cause interaction. This is a stronger condition than regular positive additive interaction which only requires $p_{11} - p_{10} - p_{01} + p_{00} > 0$ because with the condition $p_{11} - p_{10} - p_{01} > 0$ we are no longer adding back in the outcome probability p_{00} for the doubly unexposed group. The condition $p_{11} - p_{10} - p_{01} > 0$ expressed in terms of *RERI* is equivalent to *RERI* > 1.

VanderWeele (2010a b) discussed empirical tests for an even stronger notion of interaction. We might say that there is a “singular” or “epistatic” interaction if there are individuals in the population who will have the outcome if and only if both exposures are present; in counterfactual notation, that is, there are individuals for whom $Y_{11} = 1$ but $Y_{10} = Y_{01} = Y_{00} = 0$. In the genetics literature, when gene-gene interactions are considered, such response patterns are sometimes called instances of “compositional epistasis” (Phillips, 2008; Cordell, 2009) and constitute settings in which the effect of one genetic factor is masked unless the other is present. VanderWeele (2010a b) noted that if the effects of the two exposures on the outcome were unconfounded then

$$p_{11}-p_{10}-p_{01}-p_{00}>0$$

would imply the presence of such an "epistatic interaction". Again this is an even stronger notion of interaction in that we are now subtracting p_{00}. The condition $p_{11} - p_{10} - p_{01} - p_{00} > 0$ expressed in terms of *RERI* is equivalent to *RERI* > 2.

It is relatively straightforward to derive sample size and power formulae for tests for such sufficient cause or epistatic interactions. The sample size for *RERI* given above could be used but for sufficient cause interaction, to test *RERI* > 1, one would replace the η in the denominator of the sample size formula by $(\eta - 1)$; and for epistatic interaction, to test *RERI* > 2, one would replace the η in the denominator of the formula by $(\eta - 2)$.

Thus, for cohort data, to detect a sufficient cause interaction (*RERI* > 1) at significance level α with power β when the true *RERI* is $\eta = e^{\gamma 1+\gamma 2+\gamma 3} - e^{\gamma 1} - e^{\gamma 2} + 1$, the required sample size would be

$$n=\frac{(Z_{1-\alpha/2}+Z_{\beta})^2 V_{RERI}}{(\eta-1)^2}$$

where V_{RERI} is the variance of *RERI* (see Appendix 1). And likewise, the power to detect a sufficient cause interaction for a given sample size is $Power=\Phi^{-1}\left\{-Z_{1-\alpha/2}+(\eta-1)\sqrt{(n/V_{RERI})}\right\}$. Similar formulae hold for odds ratios and using case-control data under a rare outcome: once again, one simply replaces η with $(\eta - 1)$ in all relevant formulae.

Similarly, for cohort data, to detect an epistatic interaction (*RERI* > 2) at significance level α with power β when the true *RERI* is $\eta = e^{\gamma 1+\gamma 2+\gamma 3} - e^{\gamma 1} - e^{\gamma 2} + 1$, the required sample size would be

$$n=\frac{(Z_{1-\alpha/2}+Z_{\beta})^2 V_{RERI}}{(\eta-2)^2}$$

The power to detect an epistatic interaction for a given sample size is $Power=\Phi^{-1}\left\{-Z_{1-\alpha/2}+(\eta-2)\sqrt{(n/V_{RERI})}\right\}$. Similar formulae hold for odds ratios and using case-control data under a rare outcome: one simply replaces η with $(\eta - 2)$ in all relevant formulae.

Finally, it should be noted that if it can be assumed that the effects of both exposures are positive "monotonic" in the sense that the counterfactuals Y_{ge} are non-decreasing in *g* and *e* for all individuals (i.e. the exposures never have protective effects on the outcome for any individual), then the tests $p_{11} - p_{10} - p_{01} + p_{00} > 0$ and *RERI* > 0 can be used to test for

sufficient cause interaction (VanderWeele and Robins, 2007, 2008). For epistatic interactions, if the effect of at least one of the exposures is positive monotonic (Y_{ge} is nondecreasing in at least of of g and e), then $p_{11} - p_{10} - p_{01} > 0$ suffices for an epistatic interaction the tests for $RERI > 1$ could be used; if the effect of both exposures are positive monotonic then $p_{11} - p_{10} - p_{01} + p_{00} > 0$ suffices and and tests for $RERI > 0$ could be used to test for an epistatic interaction (VanderWeele, 2010a b). To interpret interaction estimates causally, or to draw conclusions about sufficient cause or epistatic interaction, control must be made for confounding for both exposures. If control for confounding is only made for one of the two exposures the interaction estimates can still often be interpreted as measures of effect heterogeneity (VanderWeele, 2009b; Vander-Weele and Knol, 2011), i.e. of how the effect of one exposure varies across strata of the other (without commenting on the effect of the second exposure itself). Sensitivity analysis techniques for interaction and effect modification (VanderWeele and Arah, 2011; VanderWeele et al., 2012) can also be useful in assessing the impact of unmeasured confounding for interaction estimates. To interpret estimates causally, measurement error in interaction analyses should also be taken into account or corrected for (Garcia Closas et al., 1998; Zhang et al., 2008; VanderWeele, 2012); such measurement error can often lead to bias and effect estimate attenuation, and will often decrease power.

Discussion

In this paper we have derived sample size and power formulae for additive interaction in a variety of scenarios. We have considered additive interaction for absolute risks in cohort data and also the use of the relative excess risk due to interaction from logistic regression using cohort or case-control data. We saw that when the main effects were both positive then the power to detect positive interaction on the additive scale was in general greater than on the multiplicative scale. We have also discussed how the sample size and power calculations for the relative excess risk due to interaction can be easily modified to provide sample size and power calculations for causal interactions corresponding to notions of synergism in the sufficient cause framework and to notions of compositional epistasis in genetics.

As is often the case with analytic formulae for sample size and power calculations, we have not considered the consequences of control for additional covariates. In settings in which these covariates are independent of the exposures (e.g. if the exposures were both randomized) then adjustment for additional covariates should increase the power of tests (Robinson and Jewell, 1991) and in such cases the sample size and power calculations in this paper could be considered conservative estimates.

The sample size and power formulae in this paper provide additional tools for researchers to utilize additive interaction in their analyses. It is hoped that these additional tools will further encourage the use of the additive scale for interaction analysis. Not only is additive interaction more relevant for public health purposes and more closely related to mechanistic interaction in the sufficient cause framework, but as we have seen, power will often be greater to detect additive interaction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Albert PS, Ratnasinghe D, Tangrea J, Wacholder S. Limitations of the case-only design for identifying gene-environment interactions. American Journal of Epidemiology. 2001; 154:687–693. [PubMed: 11590080]

Cordell HJ. Detecting gene-gene interaction that underlie human diseases. Nature Reviews Genetics. 2009; 10:392–404.

Demidenko E. Sample size and optimal design for logistic regression with binary interaction. Statistics in Medicine. 2008; 27:36–46. [PubMed: 17634969]

Foppa I, Spiegelman D. Power and sample size calculations for case-control studies of gene-environment interactions with a polytomous exposure variable. American Journal of Epidemiology. 1997; 146:596–604. [PubMed: 9326439]

Garcia-Closas M, Lubin JH. Power and sample size calculations in case-control studies of gene-environment interactions: Comments on different approaches. American Journal of Epidemiology. 1999; 149:689–692. [PubMed: 10206617]

Garcia-Closas M, Thompson WD, Robins JM. Differential misclassification and the assessment of gene-enviornment interactions. American Journal of Epidemiology. 1998; 147:426–433. [PubMed: 9525528]

Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. American Journal of Epidemiology. 2002a; 155:478–484. [PubMed: 11867360]

Gauderman WJ. Sample size requirements for matched case-control studies of gene-environment interaction. Statistics in Medicine. 2002b; 21:35–50. [PubMed: 11782049]

Greenland S. Tests for interaction in epidemiologic studies: a review and study of power. Statistics in Medicine. 1983; 2:243–251. [PubMed: 6359318]

Greenland S. Power, sample size and smallest detectable effect determination for multivariate studies. Statistics in Medicine. 1985; 4:117–127. [PubMed: 4023473]

Hernán MA. A definition of causal effect for epidemiological studies. Journal of Epidemiology and Community Health. 2004; 58:265–271. [PubMed: 15026432]

Hosmer DW, Lemeshow S. Confidence interval estimation of interaction. Epidemiology. 1992; 3:452–56. [PubMed: 1391139]

Hwang S-J, Beaty TH, Liang K-Y, Coresh J, Khoury MJ. Minimum sample size esitimation to detect gene-environment interaction in case-control designs. American Journal of Epidemiology. 1994; 140:1029–1037. [PubMed: 7985651]

Luan J, Wong M, Day N, Wareham N. Sample size determination for studies of gene-environment interaction. International Journal of Epidemiology. 2001; 30:1035–1040. [PubMed: 11689518]

Lubin JH, Gail MH. On power and sample size for studying features of the relative odds of disease. American Journal of Epidemiology. 1990; 131:552–566. [PubMed: 2301364]

Piegorsch WW, Weinberg CR, Taylor JA. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. Statistics in Medicine. 1994; 13:153–162. [PubMed: 8122051]

Phillips PC. Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. Nature Reviews Genetics. 2008; 9:855–867.

Qiu P, Moeschberger ML, Cooke GE, Goldschmidt-Clermont PJ. Sample size to test for interaction between a specific exposure and a second risk factor in a pair-matched case-control study. Statistics in Medicine. 2000; 19:923–935. [PubMed: 10750060]

Robinson L, Jewell NP. Some surprising results about covariate adjustment in logistic regression models. International Statistical Review. 1991; 59:227–240.

Rothman KJ. Causes. American Journal of Epidemiology. 1976; 104:587–592. [PubMed: 998606]

Rothman KJ, Greenland S, Walker AM. Concepts of interaction. American Journal of Epidemiology. 1980; 112:467–470. [PubMed: 7424895]

Rothman, KJ. Modern Epidemiology. 1. Little, Brown and Company; Boston, MA: 1986.

Rothman, KJ.; Greenland, S.; Lash, TL. Modern Epidemiology. 3. Vol. chapter 5. Philadelphia: Lippincott Williams and Wilkins; 2008. Concepts of interaction.

Rubin DB. Formal modes of statistical inference for causal effects. Journal of Statistical Planning and Inference. 1990; 25:279–292.

Sturmer T, Brenner H. Flexible matching strategies to increase power and efficiency to detect and estimate gene-environment interactions in case-control studies. American Journal of Epidemiology. 2002; 155:593–602. [PubMed: 11914186]

VanderWeele TJ. Sufficient cause interactions and statistical interactions. Epidemiology. 2009a; 20:6–13. [PubMed: 19234396]

VanderWeele TJ. On the distinction between interaction and effect modification. Epidemiology. 2009b; 20:863–871. [PubMed: 19806059]

VanderWeele TJ. Empirical tests for compositional epistasis. Nature Reviews Genetics. 2010a; 11:166.

VanderWeele TJ. Epistatic interactions. Statistical Applications in Genetics and Molecular Biology. 2010b; 9:Article 1, 1–22. [PubMed: 20196744]

VanderWeele TJ. Sample size and power calculations for case-only interaction studies. Epidemiology. 2011; 22:873–874. [PubMed: 21968778]

VanderWeele TJ. Inference for additive interaction under exposure misclassification. Biometrika. 2012; 99:502–508. [PubMed: 23843668]

VanderWeele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments and confounders. Epidemiology. 2011; 22:42–52. [PubMed: 21052008]

VanderWeele TJ, Knol MJ. The interpretation of subgroup analyses in randomized trials: heterogeneity versus secondary interventions. Annals of Internal Medicine. 2011; 154:680–683. [PubMed: 21576536]

VanderWeele TJ, Mukherjee B, Chen J. Sensitivity analysis for interactions under unmeasured confounding. Statistics in Medicine. 2012 in press.

VanderWeele TJ, Robins JM. The identification of synergism in the sufficient-component cause framework. Epidemiology. 2007; 18:329–339. [PubMed: 17435441]

VanderWeele TJ, Robins JM. Empirical and counterfactual conditions for sufficient cause interactions. Biometrika. 2008; 95:49–61.

Wang S, Zhao H. Sample size needed to detect gene-gene interactions using association designs. American Journal of Epidemiology. 2003; 158:899–914. [PubMed: 14585768]

Yang Q, Khoury MJ, Flanders WD. Sample size requirements in case-only designs to detect gene-environment interaction. American Journal of Epidemiology. 1997; 146:713–719. [PubMed: 9366618]

Yang Q, Khoury MJ, Sun F, Flanders WD. Case-only design to measure gene–gene interaction. Epidemiology. 1999; 10:167–170. [PubMed: 10069253]

Zhang L, Mukherjee B, Ghosh M, Gruber S, Moreno V. Accounting for error due to misclassification of exposures in case-control studies of gene-environment interaction. Statistics in Medicine. 2008; 27:2756–2783. [PubMed: 17879261]

Appendix 1. Derivations

A.1. Derivations for additive interaction with absolute risk and cohort data

For model (1),

$$P(Y=1|G=g,E=e)=\theta_0+\theta_1 g+\theta_2 e+\theta_3 ge. \quad (1)$$

the likelihood is given by

$$L(\theta_0,\theta_1,\theta_2,\theta_3)=\prod_{i=1}^{n}(\theta_0+\theta_1 g_i+\theta_2 e_i+\theta_3 g_i e_i)^{y_i}\{1-(\theta_0+\theta_1 g_i+\theta_2 e_i+\theta_3 g_i e_i)\}^{1-y_i}$$

and the log-likelihood by

$$l(\theta_0,\theta_1,\theta_2,\theta_3)=\sum_{i=1}^{n} y_i\log(\theta_0+\theta_1 g_i+\theta_2 e_i+\theta_3 g_i e_i)+\log\{1-(\theta_0,\theta_1 g_i+\theta_2 e_i+\theta_3 g_i e_i)\}(1-y_i)$$

.

The second derivative is given by

$$\frac{\partial^2 l(\theta_0,\theta_1,\theta_2,\theta_3)}{\partial(\theta_0,\theta_1,\theta_2,\theta_3)^2}=\sum_{i=1}^{n}\frac{-y_i+2y_iQ_i-{Q_i}^2}{{Q_i}^2(1-Q_i)^2}\begin{pmatrix}1 & g_i & e_i & g_ie_i\\ g_i & g_i & g_ie_i & g_ie_i\\ e_i & g_ie_i & e_i & g_ie_i\\ g_ie_i & g_ie_i & g_ie_i & g_ie_i\end{pmatrix}$$

where $Q_i = \theta_0 + \theta_1 g_i + \theta_2 e_i + \theta_3 g_i e_i$. Let $Q = \theta_0 + \theta_1 G + \theta_2 E + \theta_3 GE$. The expected information matrix is then given by

$$\begin{aligned}I&=E\left[\frac{Y-2YQ+Q^2}{Q^2(1-Q)^2}\begin{pmatrix}1 & G & E & GE\\ G & G & GE & GE\\ E & GE & E & GE\\ GE & GE & GE & GE\end{pmatrix}\right]\\ &=E\left[E\left[\frac{Y-2YQ+Q^2}{Q^2(1-Q)^2}\begin{pmatrix}1 & G & E & GE\\ G & G & GE & GE\\ E & GE & E & GE\\ GE & GE & GE & GE\end{pmatrix}|G,E\right]\right]\\ &=E\left[E\left[\frac{1}{Q(1-Q)}\begin{pmatrix}1 & G & E & GE\\ G & G & GE & GE\\ E & GE & E & GE\\ GE & GE & GE & GE\end{pmatrix}|G,E\right]\right]\end{aligned}$$

which we may write as

$$\begin{aligned}&\frac{1}{(\theta_0)(1-\theta_0)}M_1\pi_{00}+\frac{1}{(\theta_0+\theta_1)\{1-(\theta_0+\theta_1)\}}M_2\pi_{10}\\ &+\frac{1}{(\theta_0+\theta_1+\theta_2)\{1-(\theta_0+\theta_1+\theta_2)\}}M_3\pi_{01}\\ &+\frac{1}{(\theta_0+\theta_1+\theta_2+\theta_3)\{1-(\theta_0+\theta_1+\theta_2+\theta_3)\}}M_4\pi_{11}\end{aligned}$$

where

$$M_1=\begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}, M_2=\begin{pmatrix}1&1&0&0\\1&1&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$$
$$M_3=\begin{pmatrix}1&0&1&0\\0&0&0&0\\1&0&1&0\\0&0&0&0\end{pmatrix}, M_4=\begin{pmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{pmatrix}$$

If we let $L'=\frac{1}{(\theta_0)(1-\theta_0)}\pi_{00}, F'=\frac{1}{(\theta_0+\theta_1)\{1-(\theta_0+\theta_1)\}}\pi_{10}, J'=\frac{1}{(\theta_0+\theta_2)\{1-(\theta_0+\theta_{12})\}}\pi_{01}$, and $R'=\frac{1}{(\theta_0+\theta_1+\theta_2+\theta_3)\{1-(\theta_0+\theta_1+\theta_2+\theta_3)\}}\pi_{11}$ we then have

$$I=\begin{pmatrix}L'+F'+J'+R' & F'+R' & J'+R' & R'\\ F'+R' & F'+R' & R' & R'\\ J'+R' & R' & J'+R' & R'\\ R' & R' & R' & R'\end{pmatrix}.$$

The inverse of this matrix is

$$I^{-1}=\begin{pmatrix}\frac{1}{L'} & -\frac{1}{L'} & -\frac{1}{L'} & \frac{1}{L'}\\ -\frac{1}{L'} & \frac{1}{L'}+\frac{1}{F'} & \frac{1}{L'} & -\frac{1}{L'}-\frac{1}{F'}\\ -\frac{1}{L'} & \frac{1}{L'} & \frac{1}{L'}+\frac{1}{J'} & -\frac{1}{L'}-\frac{1}{J'}\\ \frac{1}{L'} & -\frac{1}{L'}-\frac{1}{F'} & -\frac{1}{L'}-\frac{1}{J'} & \frac{1}{L'}+\frac{1}{F'}+\frac{1}{J'}+\frac{1}{R'}\end{pmatrix},$$

from which it follows $V=\frac{1}{L'}+\frac{1}{F'}+\frac{1}{J'}+\frac{1}{R'}$.

A.2. Derivations for relative excess risk due to interaction from logistic regression using cohort data

Demidenko (2008) showed that for the logistic regression model (3):

$$\log it\{P(Y=1|G=g,E=e)\}=\gamma_0+\gamma_1 g+\gamma_2 e+\gamma_3 ge. \quad (3)$$

the variance-covariance matrix for the maximum likelihood estimate of (γ_0, γ_1, γ_2, γ_3) was given by

$$\begin{pmatrix}\frac{1}{L} & -\frac{1}{L} & -\frac{1}{L} & \frac{1}{L}\\ -\frac{1}{L} & \frac{1}{L}+\frac{1}{F} & \frac{1}{L} & -\frac{1}{L}-\frac{1}{F}\\ -\frac{1}{L} & \frac{1}{L} & \frac{1}{L}+\frac{1}{J} & -\frac{1}{L}-\frac{1}{J}\\ \frac{1}{L} & -\frac{1}{L}-\frac{1}{F} & -\frac{1}{L}-\frac{1}{J} & \frac{1}{L}+\frac{1}{F}+\frac{1}{J}+\frac{1}{R}\end{pmatrix},$$

where

$$L=\frac{e^{\gamma_0}}{(1+e^{\gamma_0})^2}\pi_{00}$$
$$F=\frac{e^{\gamma_0+\gamma_1}}{(1+e^{\gamma_0+\gamma_1})^2}\pi_{10}$$
$$J=\frac{e^{\gamma_0+\gamma_2}}{(1+e^{\gamma_0+\gamma_1})^2}\pi_{01}$$
$$R=\frac{e^{\gamma_0+\gamma_1+\gamma_2+\gamma_3}}{(1+e^{\gamma_0+\gamma_1+\gamma_2+\gamma_3})^2}\pi_{11}.$$

From the delta method, it follows that the variance of $\hat{RERI}=e^{\hat{\gamma}_1+\hat{\gamma}_2+\hat{\gamma}_3}-e^{\hat{\gamma}_1}-e^{\hat{\gamma}_2}+1$ is given by

$$\begin{pmatrix} 0 \\ e^{\gamma_1+\gamma_2+\gamma_3}-e^{\gamma_1} \\ e^{\gamma_1+\gamma_2+\gamma_3}-e^{\gamma_2} \\ e^{\gamma_1+\gamma_2+\gamma_3} \end{pmatrix}' \begin{pmatrix} \frac{1}{L} & -\frac{1}{L} & -\frac{1}{L} & \frac{1}{L} \\ -\frac{1}{L} & \frac{1}{L}+\frac{1}{F} & \frac{1}{L} & -\frac{1}{L}-\frac{1}{F} \\ -\frac{1}{L} & \frac{1}{L} & \frac{1}{L}+\frac{1}{J} & -\frac{1}{L}-\frac{1}{J} \\ \frac{1}{L} & -\frac{1}{L}-\frac{1}{F} & -\frac{1}{L}-\frac{1}{J} & \frac{1}{L}+\frac{1}{F}+\frac{1}{J}+\frac{1}{R} \end{pmatrix} \times \begin{pmatrix} 0 \\ e^{\gamma_1+\gamma_2+\gamma_3}-e^{\gamma_1} \\ e^{\gamma_1+\gamma_2+\gamma_3}-e^{\gamma_2} \\ e^{\gamma_1+\gamma_2+\gamma_3} \end{pmatrix}$$
$$= \begin{pmatrix} 0 \\ e^{\gamma_1+\gamma_2+\gamma_3}-e^{\gamma_1} \\ e^{\gamma_1+\gamma_2+\gamma_3}-e^{\gamma_2} \\ e^{\gamma_1+\gamma_2+\gamma_3} \end{pmatrix}' \begin{pmatrix} -\frac{1}{L}e^{\gamma_1+\gamma_2+\gamma_3}+\frac{1}{L}(e^{\gamma_1}+e^{\gamma_2} \\ \frac{1}{L}e^{\gamma_1+\gamma_2+\gamma_3}-(\frac{1}{L}+\frac{1}{F})e^{\gamma_1}-\frac{1}{L}e^{\gamma_2} \\ \frac{1}{L}e^{\gamma_1+\gamma_2+\gamma_3}-\frac{1}{L}e^{\gamma_1}-(\frac{1}{L}+\frac{1}{J})e^{\gamma_2} \\ (-\frac{1}{L}+\frac{1}{R})e^{\gamma_1+\gamma_2+\gamma_3}+(\frac{1}{L}+\frac{1}{F})e^{\gamma_1}+(\frac{1}{L}+\frac{1}{J})e^{\gamma_2} \end{pmatrix}$$
$$=e^{\gamma_1+\gamma_2+\gamma_3}\{\frac{1}{L}e^{\gamma_1+\gamma_2+\gamma_3}-(\frac{1}{L}+\frac{1}{F})e^{\gamma_1}-\frac{1}{L}e^{\gamma_2}+\frac{1}{L}e^{\gamma_1+\gamma_2+\gamma_3}-\frac{1}{L}e^{\gamma_1}$$
$$-(\frac{1}{L}+\frac{1}{J})e^{\gamma_2}+(-\frac{1}{L}+\frac{1}{R})e^{\gamma_1+\gamma_2+\gamma_3}+(\frac{1}{L}+\frac{1}{F})e^{\gamma_1}+(\frac{1}{L}+\frac{1}{J})e^{\gamma_2}\}$$
$$-e^{\gamma_1}\{\frac{1}{L}e^{\gamma_1+\gamma_2+\gamma_3}-(\frac{1}{L}+\frac{1}{F})e^{\gamma_1}+\frac{1}{L}e^{\gamma_2}\}$$
$$-e^{\gamma_2}\{\frac{1}{L}e^{\gamma_1+\gamma_2+\gamma_3}-\frac{1}{L}e^{\gamma_1}-(\frac{1}{L}+\frac{1}{J}e^{\gamma_2}\}$$
$$=(\frac{1}{L}+\frac{1}{R})e^{2(\gamma_1+\gamma_2+\gamma_3)}-\frac{2}{L}e^{2\gamma_1+\gamma_2+\gamma_3}-\frac{2}{L}e^{\gamma_1+2\gamma_2+\gamma_3}+(\frac{1}{L}+\frac{1}{F})e^{2\gamma_1}+(\frac{1}{L}+\frac{1}{J})e^{2\gamma_2}+\frac{2}{L}e^{\gamma_1+\gamma_2}.$$

A.3. Derivations for multiplicative and additive interaction for the log-linear model

For the log-linear model,

$$\log\{P(Y=1|G=g,E=e)\}=\kappa_0+\kappa_1 g+\kappa_2 e+\kappa_3 ge. \quad (5)$$

suppose we wish to use a Wald test for the null hypothesis $\kappa_3 = 0$. The sample size required to detect an multiplicative interaction of magnitude $\kappa_3 = \eta$ with significance level α and power β is

$$n=\frac{(Z_{1-\alpha/2}+Z_\beta)^2 V_{mult(RR)}}{\eta^2}$$

where $Z_{1-\alpha/2}$ and Z_β are the $(1-\alpha/2)$th and βth quantiles respectively of the standard normal distribution and where $V_{mult(RR)}$ is the variance of $\hat{\kappa_3}$ under the alternative that $\kappa_3 = \eta$. Likewise, we can calculate the power for a given sample size using

$Power = \Phi^{-1}\{-Z_{1-\alpha/2} + \eta\sqrt{(n/V_{mult(RR)})}\}$. The variance $V_{mult(RR)}$ can be derived as follows. The likelihood is given by

$$L(\kappa_0,\kappa_1,\kappa_2,\kappa_3) = \prod_{i=1}^{n} e^{(\kappa_0+\kappa_1 g_i+\kappa_2 e_i+\kappa_3 g_i e_i)y_i}\{1-e^{(\kappa_0+\kappa_1 g_i+\kappa_2 e_i+\kappa_3 g_i e_i)}\}^{1-y_i}$$

and the log-likelihood by

$$l(\kappa_0,\kappa_1,\kappa_2,\kappa_3) = \sum_{i=1}^{n} y_i(\kappa_0+\kappa_1 g_i+\kappa_2 e_i+\kappa_3 g_i e_i) + \log\{1-e^{(\kappa_0,\kappa_1 g_i+\kappa_2 e_i+\kappa_3 g_i e_i)}\}(1-y_i).$$

The second derivative is given by

$$\frac{\partial^2 l(\kappa_0,\kappa_1,\kappa_2,\kappa_3)}{\partial(\kappa_0,\kappa_1,\kappa_2,\kappa_3)^2} = \sum_{i=1}^{n} \frac{-(1-y_i)Q_i}{(1-Q_i)^2}\begin{pmatrix} 1 & g_i & e_i & g_i e_i \\ g_i & g_i & g_i e_i & g_i e_i \\ e_i & g_i e_i & e_i & g_i e_i \\ g_i e_i & g_i e_i & g_i e_i & g_i e_i \end{pmatrix}$$

where $Q_i = e^{\kappa 0+\kappa 1 g_i+\kappa 2 e_i+\kappa 3 g_i e_i}$. Let $Q = e^{\kappa 0+\kappa 1 G+\kappa 2 E+\kappa 3 GE}$. The expected information matrix is then given by

$$\begin{aligned} I &= E\left[\frac{(1-Y)Q}{(1-Q)^2}\begin{pmatrix} 1 & G & E & GE \\ G & G & GE & GE \\ E & GE & E & GE \\ GE & GE & GE & GE \end{pmatrix}\right] \\ &= E\left[E\left[\frac{(1-Y)Q}{(1-Q)^2}\begin{pmatrix} 1 & G & E & GE \\ G & G & GE & GE \\ E & GE & E & GE \\ GE & GE & GE & GE \end{pmatrix}|G,E\right]\right] \\ &= E\left[E\left[\frac{Q}{(1-Q)}\begin{pmatrix} 1 & G & E & GE \\ G & G & GE & GE \\ E & GE & E & GE \\ GE & GE & GE & GE \end{pmatrix}|G,E\right]\right] \end{aligned}$$

which we may write as

$$\frac{e^{\kappa_0}}{(1-e^{\kappa_0})}M_1\pi_{00} + \frac{e^{\kappa_0+\kappa_1}}{(1-e^{\kappa_0+\kappa_1})}M_2\pi_{10} + \frac{e^{\kappa_0+\kappa_2}}{(1-e^{\kappa_0+\kappa_2})}M_3\pi_{01} + \frac{e^{\kappa_0+\kappa_1+\kappa_2+\kappa_3}}{(1-e^{\kappa_0+\kappa_1+\kappa_2+\kappa_3})}M_4\pi_{11}$$

where

$$M_1=\begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}, M_2=\begin{pmatrix}1&1&0&0\\1&1&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$$
$$M_3=\begin{pmatrix}1&0&1&0\\0&0&0&0\\1&0&1&0\\0&0&0&0\end{pmatrix}, M_4=\begin{pmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{pmatrix}$$

If we Let $L^\dagger=\frac{e^{\kappa_0}}{(1-e^{\kappa_0})}\pi_{00}$, $F^\dagger=\frac{e^{\kappa_0+\kappa_1}}{(1-e^{\kappa_0+\kappa_1})}\pi_{10}$, $J^\dagger=\frac{e^{\kappa_0+\kappa_2}}{(1-e^{\kappa_0+\kappa_2})}\pi_{01}$, and $R^\dagger=\frac{e^{\kappa_0+\kappa_1+\kappa_2+\kappa_3}}{(1-e^{\kappa_0+\kappa_1+\kappa_2+\kappa_3})}\pi_{11}$ we then have

$$I=\begin{pmatrix}L^\dagger+F^\dagger+J^\dagger+R^\dagger & F^\dagger+R^\dagger & J^\dagger+R^\dagger & R^\dagger\\ F^\dagger+R^\dagger & F^\dagger+R^\dagger & R^\dagger & R^\dagger\\ J^\dagger+R^\dagger & R^\dagger & J^\dagger+R^\dagger & R^\dagger\\ R^\dagger & R^\dagger & R^\dagger & R^\dagger\end{pmatrix}.$$

The inverse of this matrix is

$$I^{-1}=\begin{pmatrix}\frac{1}{L^\dagger} & -\frac{1}{L^\dagger} & -\frac{1}{L^\dagger} & \frac{1}{L^\dagger}\\ -\frac{1}{L^\dagger} & \frac{1}{L^\dagger}+\frac{1}{F^\dagger} & \frac{1}{L^\dagger} & -\frac{1}{L^\dagger}-\frac{1}{F^\dagger}\\ -\frac{1}{L^\dagger} & \frac{1}{L^\dagger} & \frac{1}{L^\dagger}+\frac{1}{J^\dagger} & -\frac{1}{L^\dagger}-\frac{1}{J^\dagger}\\ \frac{1}{L^\dagger} & -\frac{1}{L^\dagger}-\frac{1}{F^\dagger} & -\frac{1}{L^\dagger}-\frac{1}{J^\dagger} & \frac{1}{L^\dagger}+\frac{1}{F^\dagger}+\frac{1}{J^\dagger}+\frac{1}{R^\dagger}\end{pmatrix},$$

from which it follows $V=\frac{1}{L^\dagger}+\frac{1}{F^\dagger}+\frac{1}{J^\dagger}+\frac{1}{R^\dagger}$.

The RERI from log-linear model (5) is given by:

$$RERI=e^{\kappa_1+\kappa_2+\kappa_3}-e^{\kappa_1}-e^{\kappa_2}+1.$$

Suppose we wish to use a Wald test for the null hypothesis *RERI* = 0. The sample size required to detect a *RERI* of magnitude $\eta = e^{\kappa 1+\kappa 2+\kappa 3}-e^{\kappa 1}-e^{\kappa 2}+1$ with significance level α and power β is

$$n=\frac{(Z_{1-\alpha/2}+Z_\beta)^2 V_{RERI(RR)}}{\eta^2}$$

where $Z_{1-\alpha/2}$ and Z_β are the $(1-\alpha/2)$th and βth quantiles respectively of the standard normal distribution and where $V_{RERI(RR)}$ is the variance of $RERI = e^{\hat{\kappa}1+\hat{\kappa}2+\hat{\kappa}3}-e^{\hat{\kappa}1}-e^{\hat{\kappa}2}+1$ under the alternative. Likewise, to calculate the power for a given sample size we could use

$$Power=\Phi^{-1}\left\{-Z_{1-\alpha/2}+\eta\sqrt{(n/V_{RERI(RR)})}\right\}.$$

Using an argument analogous to that in Appendix A.2 we have that

$$V^*_{RERI(RR)}=(\frac{1}{L^\dagger}+\frac{1}{R^\dagger})e^{2(\kappa_1+\kappa_2+\kappa_3)}-\frac{2}{L^\dagger}e^{2\kappa_1+\kappa_2+\kappa_3}-\frac{2}{L^\dagger}e^{\kappa_1+2\kappa_2+\kappa_3}+(\frac{1}{L^\dagger}+\frac{1}{F^\dagger})e^{2\kappa_1}+(\frac{1}{L^\dagger}+\frac{1}{J^\dagger})e^{2\kappa_2}+\frac{2}{L^\dagger}e^{\kappa_1+\kappa_2}.$$

A.4 Derivations for Case-Control Exposure Probabilities from the Probabilities in the Underlying Population

Here we derive the proportions in each joint exposure group in the case-control sample, $\pi^*_{00}, \pi^*_{10}, \pi^*_{01}, \pi^*_{11}$, from the proportion in each joint exposure group in the underlying population, π_{00}, π_{10}, π_{01}, π_{11}, under an assumption that the outcome is rare. We will use $P^*(\cdot)$ to denote probabilities in the case-control sample and $P(\cdot)$ to denote probabilities in the underlying population. We have that

$$\begin{aligned}\pi^*_{ge}=P^*(G=g,E=e)&=P^*(G=g,E=e|Y=0)P^*(Y=0)+P^*(G=g,E=e|Y=1)P^*(Y=1)\\&=P(G=g,E=e|Y=0)P^*(Y=0)+P(G=g,E=e|Y=1)P^*(Y=1)\\&\approx \pi_{ge}=P^*(Y=0)+P(G=g,E=e|Y=1)P^*(Y=1)\end{aligned}$$

where the final equality follows because the outcome is rare and thus the exposure distribution among the controls will approximate that in the underlying population. We then also have that

$$\begin{aligned}P(G=g,E=e|Y=1)&=\frac{P(Y=1|G=g,E=e)P(G=g,E=e)}{P(Y=1)}\\&=\frac{P(Y=1|G=g,E=e)P(G=g,E=e)}{\sum_{i,j}P(Y=1|G=i,E=j)P(G=i,E=j)}\\&=\frac{\frac{P(Y=1|G=g,E=e)}{P(Y=1|G=0,E=0)}\pi_{ge}}{\sum_{i,j}\frac{P(Y=1|G=i,E=j)}{P(Y=1|G=0,E=0)}\pi_{ge}}\\&\approx\frac{\frac{P(Y=1|G=g,E=e)/\{1-P(Y=1|G=g,E=e)\}}{P(Y=1|G=0,E=0)/\{1-P(Y=1|G=0,E=0)\}}\pi_{ge}}{\sum_{i,j}\frac{P(Y=1|G=i,E=j)/\{1-P(Y=1|G=i,E=j)\}}{P(Y=1|G=0,E=0)/\{1-P(Y=1|G=0,E=0)\}}\pi_{ij}}\end{aligned}$$

where the final equality follows from the rare outcome assumption which implies that risk ratios approximate odds ratio. The odds ratios can then be obtained from the specification of the parameters of the logistic regression model and we thus obtain:

$$\pi^*_{00} \approx \pi_{00}P^*(Y=0)+\frac{\pi_{00}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)$$
$$\pi^*_{10} \approx \pi_{10}P^*(Y=0)+\frac{e^{\gamma_1}\pi_{10}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)$$
$$\pi^*_{01} \approx \pi_{01}P^*(Y=0)+\frac{e^{\gamma_2}\pi_{01}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)$$
$$\pi^*_{11} \approx \pi_{11}P^*(Y=0)+\frac{e^{\gamma_1+\gamma_2+\gamma_3}\pi_{11}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)$$

Under this rare outcome assumption we can also obtain $\gamma_0 = \log\{P^*(Y = 1|G = 0, E = 0)/P^*(Y = 0|G = 0, E = 0)\}$, the log odds of baseline probability of the outcome in doubly unexposed group in the case-control sample, from $P^*(Y = 0)$ and $P^*(Y = 1)$ because

$$\begin{aligned}P^*(Y=1|G=0,E=0) &= \frac{P^*(G=0,E=0|Y=1)P^*(Y=1)}{P^*(G=0,E=0)} \\ &= \frac{P(G=0,E=0|Y=1)P^*(Y=1)}{P^*(G=0,E=0)} \\ &\approx \frac{\frac{\pi_{00}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)}{\pi_{00}P^*(Y=0)+\frac{\pi_{00}}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}}P^*(Y=1)} \\ &= \frac{P^*(Y=1)}{\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3}P^*(Y=0)+P^*(Y=1)} \\ &= 1/\{1+(\pi_{00}+\pi_{10}e^{\gamma_1}+\pi_{01}e^{\gamma_2}+\pi_{11}e^{\gamma_1+\gamma_2+\gamma_3})P^*(Y=0)/P^*(Y=1)\}.\end{aligned}$$

If instead the proportions in each joint exposure group in the case-control sample are specified, $\pi^*_{00}, \pi^*_{10}, \pi^*_{01}, \pi^*_{11}$, then we could obtain γ_0 by numerically solving

$$P^*(Y=0)=\frac{\pi^*_{00}}{1+e^{\gamma_0}}+\frac{\pi^*_{10}}{1+e^{\gamma_0+\gamma_1}}+\frac{\pi^*_{01}}{1+e^{\gamma_0+\gamma_2}}+\frac{\pi^*_{11}}{1+e^{\gamma_0+\gamma_1+\gamma_2+\gamma_3}}$$

for γ_0. If the joint or marginal exposure probabilities are specified separately for the cases and controls then under an assumption of a rare outcome, the distribution of the exposures amongst the controls could be used as an approximation to π_{00}, π_{10}, π_{01}, π_{11} or π_g, π_e, Δ.

Appendix 2. Epidemiologic Practice: Excel Spreadsheets for Sample Size and Power Calculations for Additive and Multiplicative Interaction

As part of the online supplement for this paper there are two Excel spread-sheets that will automatically perform power and sample size calculations for additive and multiplicative interaction for (i) cohort and (ii) case-control, and case-only data. All of these spreadsheets return sample size and power calculations for the Wald test statistic for additive or multiplicative interaction with variance calculated under the alternative (cf. Demidenko, 2008; VanderWeele, 2011).

The first spreadsheet performs power and sample size calculations for additive and multiplicative interaction for cohort data. For the power calculations, the user has the option of entering marginal exposure probabilities and the odds ratio relating the prevalence of both exposures (Sheet 1) or the joint exposure probabilities (Sheet 2). On Sheet 1, the user inputs the significance level of the test (alpha), the sample size (n), the probability of the outcome in the doubly unexposed reference group (p00), the main effect odds ratio for the first

exposure (OR10), the main effect odds ratio for the second exposure (OR01), the odds ratio multiplicative interaction (IOR=$OR_{11}/(OR_{10}*OR_{01})$), the marginal prevalence of the first exposure (P(G=1)), the marginal prevalence of the second exposure (P(E=1)) and the odds ratio relating the dependence between the two exposures (OR_GE). The Excel spreadsheet returns both one-sided power (to detect positive interaction) and two-sided power (to detect positive or negative interaction) for (i) additive interaction on the risk difference scale, (ii) multiplicative interaction on the risk ratio scale, (iii) multiplicative interaction on the odds ratio scale, (iv) additive interaction using the relative excess risk due to interaction (RERI; cf. Hosmer and Lemeshow, 1992) for risk ratios, and (v) additive interaction using the relative excess risk due to interaction for odds ratios, assuming a rare outcome. On Sheet 2, the user specifies the same inputs except that instead of the marginal probabilities and odds ratio relating the exposures (P(G=1), P(E=1), OR_GE), the user specifies the joint exposure probabilities for each of the four possible exposure combinations (in the Excel spreadsheet these are pi00, pi10, pi01, pi11). The Excel spreadsheet then again returns items (i)–(v) above.

For sample size calculations from cohort data, the user has the option of entering marginal exposure probabilities and the odds ratio relating the prevalence of both exposures (Sheet 3) or the joint exposure probabilities (Sheet 4). The user specifies exactly the same parameters as the spreadsheet for power calculations for cohort data except that instead of specifying the sample size, the power is specified (Power), and the Excel spreadsheet returns the required sample size for a test of the specified significance level and power to detect (i) additive interaction on the risk difference scale, (ii) multiplicative interaction on the risk ratio scale, (iii) multiplicative interaction on the odds ratio scale, (iv) additive interaction using the relative excess risk due to interaction (RERI) for risk ratios, (v) additive interaction using the relative excess risk due to interaction for odds ratios, assuming a rare outcome.

The second spreadsheet performs power and sample size calculations for additive and multiplicative interaction for case-control and case-only data. For power calculations (Sheet 1), the user inputs the significance level of the test (alpha), the number of cases (n Cases) and number of controls (n Controls), the main effect odds ratio for the first exposure (OR10), the main effect odds ratio for the second exposure (OR01), the odds ratio multiplicative interaction (IOR), the marginal prevalence of the first exposure (P(G=1)), the marginal prevalence of the second exposure (P(E=1)) and the odds ratio relating the dependence between the two exposures (OR_GE). The Excel spreadsheet returns both one-sided power (to detect positive interaction) and two-sided power (to detect positive or negative interaction) for (i) additive interaction using the relative excess risk due to interaction (RERI) for odds ratios and (ii) multiplicative interaction on the odds ratio scale. If the two exposures are specified as independent (i.e. if OR_GE is specified as 1) then the spreadsheet will also return the power for the case-only estimator of multiplicative interaction (cf. Piegorsch et al, 1994; Yang et al., 1999) based on the number of cases. If the two exposures are not specified as independent (i.e. if OR_GE is specified as any number other than 1), the spreadsheet will return “#DIV/0!” for the power for the case-only estimator indicating that the case-only test is inapplicable in this setting because the two

exposures are not independent. All power calculations for the case-control and case-only power spreadsheet make a rare outcome assumption. The power calculations are based on the variance calculated under the alternative (as in Demidenko (2008) for logistic regression multiplicative interactions and VanderWeele (2011) for case-only multiplicative interactions) rather the variance calculated under the null, as the variance under the alternative corresponds to the test statistics that are commonly used in practice.

For sample size calculations for additive and multiplicative interaction for case-control and case-only data (Sheet 2), the user inputs the significance level of the test (alpha), the proportion of cases in the case-control sample (Cs/(Cs+Cont)), the desired power of the test (Power), the main effect odds ratio for the first exposure (OR10), the main effect odds ratio for the second exposure (OR01), the odds ratio multiplicative interaction (IOR), the marginal prevalence of the first exposure (P(G=1)), the marginal prevalence of the second exposure (P(E=1)) and the odds ratio relating the dependence between the two exposures (OR_GE). The Excel spreadsheet returns the required sample size for a test of the specified significance level and power for (i) additive interaction using the relative excess risk due to interaction (RERI) for odds ratios and (ii) multiplicative interaction on the odds ratio scale. If the two exposures are specified as independent (i.e. if OR_GE is specified as 1) then the spreadsheet will also return the required sample size, i.e. number of cases, to detect multiplicative interaction for the case-only estimator of multiplicative interaction. If the two exposures are not specified as independent (i.e. if OR_GE is specified as any number other than 1), the spreadsheet will return “#DIV/0!” for the required sample size for the case-only estimator indicating that the case-only test is inapplicable in this setting because the two exposures are not independent. All power calculations for the case-control and case-only sample size spreadsheet make a rare outcome assumption. The sample size calculations are based on the variance calculated under the alternative as this corresponds to the test statistics that are commonly used in practice (cf. Garcia-Closas and Lubin, 1999; Demidenko, 2008; VanderWeele, 2011).

Table 1

Power to detect additive interaction and multiplicative interaction for various sample sizes, main effects, and interaction parameters (first number in each column is power to detect additive interaction; second number is power for multiplicative interaction)

I_{OR}	OR_{10}	OR_{01}	n = 500	n = 1000	n = 3000	n = 5000
1.1	1	1	.05, .05	.06, .06	.10, .09	.14, .13
1.1	1.3	1.3	.07, .04	.10, .05	.23, .09	.34, .12
1.1	1.5	1.8	.13, .04	.23, .05	.55, .08	.77, .11
1.3	1	1	.12, .11	.21, .17	.50, .42	.72, .62
1.3	1.3	1.3	.18, .10	.32, .15	.73, .37	.91, .56
1.3	1.5	1.8	.27, .09	.48, .14	.91, .33	.99, .50
1.5	1	1	.25, .19	.44, .34	.88, .77	.98, .93
1.5	1.3	1.3	.32, .17	.56, .30	.95, .70	1.00, .89
1.5	1.5	1.8	.40, .15	.68, .26	.99, .63	1.00, .84
2	1	1	.57, .44	.85, .73	1.00, .99	1.00, 1.00
2	1.3	1.3	.58, .39	.86, .65	1.00, .98	1.00, 1.00
2	1.5	1.8	.59, .34	.87, .59	1.00, .97	1.00, 1.00
3	1	1	.81, .77	.98, .97	1.00, 1.00	1.00, 1.00
3	1.2	1.3	.74, .70	.96, .94	1.00, 1.00	1.00, 1.00
3	1.5	1.8	.68, .62	.93, .89	1.00, 1.00	1.00, 1.00