

Ola Hössjer\*, Ingrid Kockum, Lars Alfredsson, Anna Karin Hedström, Tomas Olsson and Magnus Lekman

# A General Framework for and New Normalization of Attributable Proportion

DOI 10.1515/em-2015-0028

**Abstract:** A unified theory is developed for attributable proportion (AP) and population attributable fraction (PAF) of joint effects, marginal effects or interaction among factors. We use a novel normalization with a range between  $-1$  and  $1$  that gives the traditional definitions of AP or PAF when positive, but is different when they are negative. We also allow for an arbitrary number of factors, both those of primary interest and confounders, and quantify interaction as departure from a given model, such as a multiplicative, additive odds or disjunctive one. In particular, this makes it possible to compare different types of threeway or higher order interactions. Effect parameters are estimated on a linear or logit scale in order to find point estimates and confidence intervals for the various versions of AP and PAF, for prospective or retrospective studies. We investigate the accuracy of three confidence intervals; two of which use the delta method and a third bootstrapped interval. It is found that the delta method with logit type transformations, and the bootstrap, perform well for a wide range of models. The methodology is also applied to a multiple sclerosis (MS) data set, with smoking and two genetic variables as risk factors.

**Keywords:** attributable proportion, confidence interval, models of no interaction, normalization, population attributable fraction

## 1 Introduction

Most complex diseases are influenced by a number of genetic and environmental factors, co-factors or predictors that either have a marginal effect and/or interact in some way (Maher 2008; Eichler et al. 2010; Bookman et al. 2011). The impact of these factors on disease risk is specified with a penetrance function

$$\theta(x) = P(Y = 1 | X = x), \quad (1)$$

where  $Y$  equals 1 (0) for an individual with (without) the disease, and  $x = (x_1, \dots, x_p)$  is a vector with exposure status for  $p$  co-factors. We will assume that the factors are binary, with  $x_i = 1$  or 0 depending on whether an individual is exposed to factor  $i$  or not, although extensions to continuous covariates will be discussed in Section 8. Often a few of the variables  $J \subset \{1, \dots, p\}$  are of main interest, whereas the others are confounders. In this paper we develop a unified theory for the attributable risk or attributable proportion (AP), and the population attributable fraction (PAF), when the effects of confounders are taken into account. This theory will involve both joint effects, marginal effects and interaction.

The joint effect of all non-confounders will refer to their cumulative marginal effects and interaction. The traditional definition of AP;

$$AP_{\text{trad}}(x) = \frac{\theta(x) - \theta_{\text{rem}}(x)}{\theta(x)}, \quad (2)$$

\*Corresponding author: Ola Hössjer, Department of Mathematics, Stockholm University, Stockholm, Sweden, E-mail: ola@math.su.se

Ingrid Kockum, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Lars Alfredsson, Anna Karin Hedström, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

Tomas Olsson, Magnus Lekman, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

is the joint fraction of disease risk due the factors in  $J$ . The penetrance function  $\theta_{\text{rem}}(x)$  is the disease risk of an individual whose exposure to all factors  $i \in J$  of interest has been *removed*, whereas the confounding variables  $i \notin J$  remain at the levels specified by  $x$ . Equation (2) generalizes the definition of attributable proportion, first introduced by Levin (1953) to quantify the effect that one single risk factor, smoking, has on lung cancer. The joint effect is then equivalent to a marginal effect, since it is not possible to have interaction for one single factor. Greenland and Robins (1988) noted that sometimes there is an ambiguity as to whether  $\theta_{\text{rem}}$  should include all subjects that develop the disease without exposure to the relevant factor(s), or only those for which the factor(s) has no causal role. Often the former definition of attributable proportion is used, since it is easier to estimate from data.

Since smoking is a risk factor ( $\theta > \theta_{\text{rem}}$ ),  $\text{AP}_{\text{trad}}$  is a number between 0 and 1. On the other hand, it is negative without any lower bound for preventive factors ( $\theta < \theta_{\text{rem}}$ ). It is then possible to use the preventive fraction

$$\text{PF}(x) = \frac{\theta(x) - \theta_{\text{rem}}(x)}{\theta_{\text{rem}}(x)}, \quad (3)$$

a number between 0 and 1 that equals the fraction of reduced disease risk for a subject exposed to preventive factors, see for instance Aschengrau et al. (2003, pp.65–66). But it is redundant to use two different quantities  $\text{AP}_{\text{trad}}$  and PF in order to quantify the effect of one set of factors, in particular when it is not known before a study whether a predictor increases or decreases risk. The relative risk could replace eqs (2) and (3) but it has no upper bound for risk factors. Instead, we propose another normalization

$$\text{AP}(x) = \frac{\theta(x) - \theta_{\text{rem}}(x)}{\max[\theta(x), \theta_{\text{rem}}(x)]} \quad (4)$$

of the attributable proportion, a number between  $-1$  and  $1$  that equals  $\text{AP}_{\text{trad}}$  for risk factors and  $-\text{PF}$  for preventive factors.

In order to define a version of AP that averages the effect of all confounders, while still controlling for their effect, we introduce  $\theta_{\text{add}}(x)$ , the penetrance function for a subject whose exposure to all factors  $i \in J$  has been *added* or turned on, whereas all confounding factors  $i \notin J$  remain at the levels specified by  $x$ . Let  $\mathbf{x} = (X_1, \dots, X_p)$  refer to the exposure vector of a randomly chosen subject in the population. The joint proportion of risk due to the predictors in  $J$  is

$$\text{AP} = \frac{E[\theta_{\text{add}}(X)] - E[\theta_{\text{rem}}(X)]}{\max\{E[\theta_{\text{add}}(X)], E[\theta_{\text{rem}}(X)]\}}, \quad (5)$$

when averaging the effect of all confounders. The numerator of eq. (5) is the average causal effect (ACE) when causation can be replaced by association (Agresti 2013), and the denominator normalizes ACE to a value between  $-1$  and  $1$ .

It is also possible to define counterparts of eqs (2), (3) and (4) for a whole population. The population attributable fraction

$$\text{PAF}_{\text{trad}} = \frac{E[\theta(X)] - E[\theta_{\text{rem}}(X)]}{E[\theta(X)]}, \quad (6)$$

quantifies the fraction of disease prevalence  $E[\theta(X)]$  that is jointly attributable to a subset  $J$  of risk factors. It is an important quantity for public health and a tool for deciding which possible interventions that should be prioritized, see Deubner et al. (1980), Northridge (1995), Benichou (2001) and Sjölander and Vansteelandt (2011). For preventive factors, one uses instead the population prevented fraction

$$\text{PPF} = \frac{E[\theta(X)] - E[\theta_{\text{rem}}(X)]}{E[\theta_{\text{rem}}(X)]}, \quad (7)$$

i. e. the fraction by which prevalence is decreased when the population is jointly exposed to all factors in  $J$ . Analogously to eq. (4), we combine eqs (6) and (7) into a generalized version

$$\text{PAF} = \frac{E[\theta(X)] - E[\theta_{\text{rem}}(X)]}{\max\{E[\theta(X)], E[\theta_{\text{rem}}(X)]\}} \quad (8)$$

of population attributable fraction that takes values between  $-1$  and  $1$  and equals  $\text{PAF}_{\text{trad}}$  for risk factors and  $-\text{PPF}$  for preventive factors.

It is possible to reinterpret eqs (4), (5) and (8) to quantify epistasis or interaction among a subset  $J$  of at least two predictors, rather than their joint effect. Several methods are available to quantify interaction, see Greenland (1983), Cordell (2002), Phillips (2008), VanderWeele (2010) and VanderWeele and Knol (2014) for reviews. Some epistasis definitions involve causation, counterfactuals or potential outcomes (Rothman 1976a), and under certain conditions they can be tested statistically (VanderWeele and Robins 2008; Sjölander et al. 2014; Lekman et al. 2014). We will adopt a common and weaker definition of epistasis in terms of a model whose interaction parameter between a subset  $J$  of factors is nonzero. When attributable proportion is used to quantify the strength of this interaction (rather than the joint effect), we must not include any marginal effects. This amounts to reinterpreting  $\theta_{\text{rem}}(x)$  as the disease risk when exposure to each factor in  $J$  is the same as in the definition of  $\theta(x)$ , but interaction among them, not their marginal effects, is *removed*. This requires that a model  $M$  of no interaction is defined. Statisticians have commonly used logistic regression models and multiplicative odds, but additive models of interaction have recently gained popularity, since they are believed to approximate an underlying biological reality more accurately (Rothman 2012). In particular, a collection of factors exhibits sufficient cause interaction when they are all present in at least one causal mechanism that is sufficient to bring about the disease. For binary factors and outcomes, it has been shown in VanderWeele (2009) and references therein, that under mild regularity conditions, sufficient cause interaction is closely related to additive interaction.

Given that a model  $M$  has been specified, interaction is positive or negative when there is excess or lowered risk due to interaction, depending on the sign of  $\text{AP}_{\text{trad}}$ . For positive interaction that increases disease risk (synergism),  $\text{AP}_{\text{trad}}$  is a number between  $0$  and  $1$  that quantifies the proportion by which disease risk is lowered when interaction among the factors in  $J$  is *removed* for subjects with exposure  $x$ . But for negative interaction that lowers disease risk (antagonism),  $\text{AP}_{\text{trad}}$  has no lower bound. The alternative definition eq. (4) is identical to eq. (2) when positive, but normalized differently for negative values, as minus the proportion by which disease risk is lowered for a subject with exposure  $x$  when interaction among the factors in  $J$  is *added*, with a lower bound of  $-1$ . The two extreme cases when  $\text{AP}$  equals  $-1$  and  $1$  correspond to interaction that completely eliminates or dominates disease risk. The quantities  $\text{AP}$  and  $\text{PAF}$  in eqs (5) and (8) are also numbers between  $-1$  and  $1$ , with similar interpretation as the proportion of disease risk when interaction among the predictors in  $J$  is removed or added.

As far as we know,  $\text{AP}_{\text{trad}}$  has only been used for interaction between pairs of binary factors under additive models of no interaction, but we will apply our new normalization of  $\text{AP}$  and  $\text{PAF}$  much more generally. Our framework is a model with an arbitrary number  $p$  of factors, when the joint effect or interaction among a subset  $J$  of them is of interest. We consider a large class of models  $M$  of no interaction, such as an additive, multiplicative or disjunctive one. Results of Hosmer and Lemeshow (1992), Assman et al. (1996) and Andersson et al. (2005) are extended, whereby linear or logit parametrization of effect parameters is used to obtain point estimates and three types of confidence intervals, two Wald intervals based on the multivariate delta method and asymptotic normality of parameter estimates, and a bootstrap interval. In particular, we utilize that eq. (4) is restricted to  $[-1, 1]$  and show that a Wald interval based on a logit type transformation and the bootstrap work very well for a wide range of models.

Our work was motivated by a recent study in Lekman et al. (2014), where possible genetic interaction effects of major depressive disorder were sought for in terms of pairs of single nucleotide polymorphisms (SNPs), using different multiplicative and additive models of no interaction. For some pairs of SNPs,  $\text{AP}$

estimates for the additive models were far below  $-1$ , with difficulty of interpretation. Although different normalizations of AP due to interaction have been proposed (VanderWeele 2013), the one in eq. (4) is the first that guarantees a range of  $[-1, 1]$ .

In order to estimate AP and PAF one typically uses cohort data, and external information is needed in order to calculate these quantities from case-control data. But since odds ratios can be estimated consistently from case-control data, an alternative strategy is to redefine the attributable proportion in terms of odds ratios rather than disease risks. We apply such an odds ratio based definition of AP, with our new normalization, to estimate interaction effects of a multiple sclerosis (MS) data set previously analyzed in Hedström et al. (2011).

The paper is organized as follows. In Section 2 we consider the case of two factors and additive models of no interaction, and then in Section 3 an arbitrary number of factors, and a generalized linear model (McCullagh and Nelder 1989) family of no interaction models that include additive, additive odds, multiplicative and disjunctive models. In Section 4 we define the analogous versions of AP in terms of odds ratios. Then we define point estimates and confidence intervals for prospective and retrospective studies in Section 5, conduct a simulation study in Section 6, analyze the MS data set in Section 7, and conclude with a discussion in Section 8. Further mathematical and implementation details are given in Appendices A-F.

## 2 Two predictors

### 2.1 Marginal and joint effects

Suppose we have  $p=2$  binary co-factors or predictors, and let

$$RR(x_1, x_2) = \frac{\theta(x_1, x_2)}{\theta(0, 0)}$$

be the relative risk for exposure profile  $(x_1, x_2)$  compared to a baseline risk when exposure is absent. The attributable proportion  $AP(x) = AP(x; J)$  in eq. (4) for a subset  $J \subset \{1, 2\}$  of predictors takes the form

$$AP(x_1, x_2; \{1\}) = \frac{\theta(x_1, x_2) - \theta(0, x_2)}{\max[\theta(x_1, x_2), \theta(0, x_2)]} = \frac{RR(x_1, x_2) - RR(0, x_2)}{\max[RR(x_1, x_2), RR(0, x_2)]} \quad (9)$$

for the marginal effect of the first factor,

$$AP(x_1, x_2; \{2\}) = \frac{\theta(x_1, x_2) - \theta(x_1, 0)}{\max[\theta(x_1, x_2), \theta(x_1, 0)]} = \frac{RR(x_1, x_2) - RR(x_1, 0)}{\max[RR(x_1, x_2), RR(x_1, 0)]} \quad (10)$$

for the marginal effect of the second factor, and

$$AP(x_1, x_2; \{1, 2\}) = \frac{\theta(x_1, x_2) - \theta(0, 0)}{\max[\theta(x_1, x_2), \theta(0, 0)]} = \frac{RR(x_1, x_2) - 1}{\max[RR(x_1, x_2), 1]} \quad (11)$$

for the joint effect of both factors. We see from eqs (9)–(11) that  $AP(x_1, x_2; J)$  is only a function of relative risk  $RR(x_1, x_2)$  when both factors are considered jointly, or when the effect of one factor is of interest and exposure from the other confounding factor is absent. In the latter case, if the confounding factor 2 (say) is exposed as well,  $AP(x_1, 1; \{1\})$  is also a function of  $RR(0, 1)$ .

The marginal effect  $AP = AP(J)$  in eq. (5) of predictor 1 is

$$AP(\{1\}) = \frac{E[\theta(1, X_2)] - E[\theta(0, X_2)]}{\max\{E[\theta(1, X_2)], E[\theta(0, X_2)]\}}.$$

where expectation is over the exposure distribution of the confounding variable  $X_2$ . The marginal effect  $AP(\{2\})$  of predictor 2 is analogous.

The population attributable fraction  $PAF = PAF(J)$  in eq. (8) is defined as an average of eqs (9)–(11) with respect to an exposure distribution  $q(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$  of both predictors. For instance, the marginal effect of predictor 1 in the population is

$$PAF(\{1\}) = \frac{E[\theta(X_1, X_2)] - E[\theta(0, X_2)]}{\max\{E[\theta(X_1, X_2)], E[\theta(0, X_2)]\}}.$$

The population attributable fraction of predictor 2 is defined in the same way, reversing the roles of  $X_1$  and  $X_2$ . The joint effect of both predictor variables is

$$PAF(\{1, 2\}) = \frac{E[\theta(X_1, X_2)] - \theta(0, 0)}{\max\{E[\theta(X_1, X_2)], \theta(0, 0)\}}.$$

## 2.2 Interaction

Both variables  $J = \{1, 2\}$  are required to define  $AP(x) = AP(x; J, M)$  of interaction, and therefore  $J$  will be dropped in the notation. We also need a model  $M$  of no interaction, and write  $\theta_{\text{rem}}(x) = \theta_{\text{rem}}(x; M)$ . The departure from no interaction is referred to as  $ERI(x) = \theta(x) - \theta_{\text{rem}}(x)$ , the excess risk due to interaction. It is usually required that both factors  $i$  are turned on ( $x_i = 1$ ) in order for  $ERI$  to be nonzero, and then it can either be positive or negative. For an additive model of no interaction the traditional definition is

$$\theta_{\text{rem,trad}}(1, 1; \text{Additive}) = \theta(1, 0) + \theta(0, 1) - \theta(0, 0).$$

Insertion into eq. (2) and division by the baseline risk  $\theta(0, 0)$  gives

$$AP_{\text{trad}}(1, 1; \text{Additive}) = \frac{RR(1, 1) - [RR(0, 1) + RR(1, 0) - 1]}{RR(1, 1)}. \quad (12)$$

The numerator of eq. (12) is commonly referred to as the interaction contrast ratio (ICR) or the relative excess risk due to interaction (RERI), whereas the denominator is the relative risk for a subject exposed to both factors. Their ratio in eq. (12) is a number between 0 and 1 that quantifies the proportion of risk due to interaction when RERI is positive, but it is negative without a lower bound when  $RERI < 0$ . In order to overcome this deficiency of  $AP_{\text{trad}}$ , let  $\theta_i(x) = P(Y = 1 | X_i = x)$  be the marginal penetrance function of factor  $i$ . It has been suggested by Rothman (2012) to recode any factor  $i$  that is preventive ( $\theta_i(1) < \theta_i(0)$ ) so that both factors become risk factors ( $\theta_i(1) > \theta_i(0)$ ). Another suggestion of Knol et al. (2011) is to recode variables jointly, so that the penetrance function is minimized by  $\theta(0, 0)$ . A third possibility is to redefine the denominator of eq. (2) to  $\theta(1, 1) - \theta(0, 0)$ , or the one in eq. (12) to  $RR(1, 1) - 1$ , see Rothman (2012) and VanderWeele (2013).

Although all three of these suggestions can be helpful, the following example shows that none of them guarantees an  $AP$  between  $-1$  and  $1$ . We assume  $\theta(0, 0) = 0$ ,  $\theta(1, 0) = \theta(0, 1) = 0.5$ ,  $\theta(1, 1) = 0.1$  and that all exposure profiles are equally likely, i. e.  $q(x_1, x_2) = 0.25$ . It follows that  $\theta_i(0) = 0.25$  and  $\theta_i(1) = 0.3$ , so that both variables are risk factors. Neither marginal nor joint risk considerations will therefore cause any recoding, but yet

$$AP_{\text{trad}}(1, 1; \text{Additive}) = \frac{0.1 - (0.5 + 0.5 - 0)}{0.1} = -9,$$

the same value as when  $\theta(1, 1) - \theta(0, 0)$  is used in the denominator instead.

Our solution is to modify eq. (12) in two ways. First we truncate the predicted risk to a number between 0 and 1, i. e.

$$\theta_{\text{rem}}(x_1, x_2) = [\theta(0, 0) + x_1\{\theta(1, 0) - \theta(0, 0)\} + x_2\{\theta(0, 1) - \theta(0, 0)\}]_0^1, \quad (13)$$

where  $[\theta]_a^b = \max[a, \min(\theta, b)]$ , arguing that it makes little sense to have a predictive risk that is not a probability. Second we apply the new normalization eq. (4) and divide by the baseline risk  $\theta(0, 0)$ , so that

$$\text{AP}(x_1, x_2; \text{Additive}) = \frac{\text{RR}(x_1, x_2) - \text{RR}_{\text{rem}}(x_1, x_2)}{\max[\text{RR}(x_1, x_2), \text{RR}_{\text{rem}}(x_1, x_2)]}, \quad (14)$$

where  $\text{RR}_{\text{rem}}(x_1, x_2) = \theta_{\text{rem}}(x_1, x_2)/\theta(0, 0)$  is the relative risk when interaction between the two factors has been removed. Defined in this way, eq. (14) is the proportion by which positive/antagonistic interaction increases risk, and minus the proportion by which protective/negative interaction decreases risk. In particular, eq. (14) equals  $-0.9$  rather than  $-9$  in the numerical example above. In order to motivate truncation in eq. (13), consider a second example with  $\theta(0, 0) = 0.3$  and  $\theta(0, 1) = \theta(1, 0) = \theta(1, 1) = 0.1$ . Then  $\text{AP}(1, 1; \text{Additive}) = 1$ , but it would equal 2 without truncation.

The excess risk due to interaction in eq. (14) is only of interest when  $x_1 = x_2 = 1$ , since otherwise it is zero. Since there are no confounders in the model ( $|J| = p = 2$ ), the attributable fraction [14] agrees with AP in eq. (5) when  $x_1 = x_2 = 1$ . But we need to consider arbitrary exposure profiles  $(x_1, x_2)$  in order to define the population counterpart  $\text{PAF} = \text{PAF}(M)$  of eq. (14) in eq. (8). For two predictors it can be written as

$$\text{PAF}(\text{Additive}) = \frac{E[\text{RR}(X_1, X_2)] - E[\text{RR}_{\text{rem}}(X_1, X_2)]}{\max\{E[\text{RR}(X_1, X_2)], E[\text{RR}_{\text{rem}}(X_1, X_2)]\}}. \quad (15)$$

When positive, eq. (15) is the fraction of population risk due to additive interaction; when negative, it is minus the fraction by which such interaction decreases risk.

### 3 Arbitrary number of predictors

#### 3.1 Marginal and joint effects

For an arbitrary number  $p$  of factors, we express in Appendix A.1 several quantities in terms of the relative risk

$$\text{RR}(x) = \frac{\theta(x)}{\theta(0)} \quad (16)$$

of exposure  $x$  compared to a scenario  $0 = (0, \dots, 0)$  where exposure is absent for all factors. This includes the joint attributable proportion  $\text{AP}(x) = \text{AP}(x; J)$  in eq. (4) for exposure  $x$ , due to factors in  $J$ , the corresponding joint attributable proportion  $\text{AP} = \text{AP}(J)$  in eq. (5), when averaging over all confounders, and the joint population attributable fraction  $\text{PAF} = \text{PAF}(J)$  for factors  $J$  in eq. (8). The marginal attributable proportion or population attributable fraction for a single non-confounding variable is a special case of these definitions, with  $|J| = 1$ .

#### 3.2 Interaction

Consider an arbitrary subset  $J$  of at least two factors, and a model  $M$  that belongs to a class of generalized linear models of no interaction. To each  $x$  we associate another vector  $x_J = (x_{J1}, \dots, x_{Jp})$ , whose exposure to all factors of interest has been *removed* ( $x_{ji} = 0$  if  $i \in J$ ), whereas the confounding variables remain at the levels specified by  $x$  ( $x_{ji} = x_i$  if  $i \notin J$ ). With this notation, we can specify the disease risk for a subject with exposure  $x$  and no interaction among the variables in  $J$ , as

$$\theta_{\text{rem}} = \theta_{\text{rem}}(x; J, M) = g^{-1} \left\{ \rho \left[ g(\theta(x_J)) + \sum_{i \in J} \beta_i x_i \right] \right\}, \quad (17)$$



where  $g$  is a link function that transforms the disease probability  $\theta$  to regression space,  $\beta_i = g[\theta(x_j + e_i)] - g[\theta(x_j)]$  is a regression coefficient for predictor variable  $i \in J$  that is typically a function of  $x_j$ ,  $e_i$  is a unit vector of length  $p$  with one in position  $i$  and zeros elsewhere, and  $\rho$  a projection that guarantees  $\theta_{\text{rem}}(x; J, M) \in [0, 1]$ .

A number of models (17) are shown in Table 1; the identity link  $g(\theta) = \theta$  gives an additive model, the odds link  $g(\theta) = \theta/(1 - \theta)$  an additive odds model, the canonical logit link  $g(\theta) = \log[\theta/(1 - \theta)]$  a multiplicative model, the log link  $g(\theta) = \log(\theta)$  a multiplicative risk model and  $g(\theta) = -\log(1 - \theta)$  a multicausal disjunctive model (Pearl 1988). No truncation is necessary ( $\rho(\eta) = \eta$ ) for the multiplicative model in order to ensure  $\theta_{\text{rem}} \in [0, 1]$ , for the additive model  $\rho(\eta) = [\eta]_0^1$ , for the multiplicative risk model  $\rho(\eta) = [\eta]_{-\infty}^0$ , and for the additive odds and disjunctive models  $\rho(\eta) = [\eta]_0^\infty$ . Our framework also incorporates many other no interaction models than eq. (17), see Appendix E.

**Table 1:** Examples of generalized linear models [17] of no interaction.

Model	Link function $g(\theta)$	$\theta(x_j)$	$\theta(x_j + e_i)$	$d(O)$	$O(x_j)$	$O(x_j + e_i)$
Additive	$\theta$	$\beta_0$	$\beta_0 + \beta_i$	$\frac{O}{1+O}$	$\frac{\beta_0}{1-\beta_0}$	$\frac{\beta_0 + \beta_i}{1-\beta_0 - \beta_i}$
Additive odds	$\frac{\theta}{1-\theta}$	$\frac{\beta_0}{1+\beta_0}$	$\frac{\beta_0 + \beta_i}{1+\beta_0 + \beta_i}$	$O$	$\beta_0$	$\beta_0 + \beta_i$
Multiplicative	$\log(\frac{\theta}{1-\theta})$	$\frac{e^{\beta_0}}{1+e^{\beta_0}}$	$\frac{e^{\beta_0 + \beta_i}}{1+e^{\beta_0 + \beta_i}}$	$\log(O)$	$e^{\beta_0}$	$e^{\beta_0 + \beta_i}$
Multipl. risk	$\log(\theta)$	$e^{\beta_0}$	$e^{\beta_0 + \beta_i}$	$\log(\frac{O}{1+O})$	$\frac{e^{\beta_0}}{1-e^{\beta_0}}$	$\frac{e^{\beta_0 + \beta_i}}{1-e^{\beta_0 + \beta_i}}$
Disjunctive	$-\log(1-\theta)$	$1 - e^{-\beta_0}$	$1 - e^{-\beta_0 - \beta_i}$	$\log(1+O)$	$e^{\beta_0} - 1$	$e^{\beta_0 + \beta_i} - 1$

Note: For each model,  $g$  is the link function that transforms the disease probability  $\theta$  to regression space  $\eta = g(\theta)$ , and  $d$  is the function that similarly links odds  $O = \theta/(1 - \theta)$  to  $\eta = d(O)$ . Risks and odds formulas are shown when exposure is turned on for at most one of the factors in  $J \subset \{1, \dots, p\}$ .

In Appendix A.2 we express a number of quantities in terms of eq. (16) and

$$\text{RR}_{\text{rem}}(x; J, M) = \frac{\theta_{\text{rem}}(x; J, M)}{\theta(0)}, \quad (18)$$

the relative risk for exposure  $x$  when interaction among all factors in  $J$  has been removed. This includes the attributable proportion  $\text{AP}(x) = \text{AP}(x; J, M)$  of interaction in eq. (4) among the factors in  $J$  for exposure  $x$ , the corresponding attributable proportion  $\text{AP} = \text{AP}(J; M)$  of interaction in eq. (5) when averaging over confounders, and, finally, the attributable fraction of interaction  $\text{PAF} = \text{PAF}(J, M)$  in eq. (8) among the factors in  $J$  for the whole population.

## 4 Odds ratios

It is possible to estimate relative risks consistently for cohort studies or case-control studies with incident cases who developed the disease after exposure (Miettinen 1976). But many case-control studies involve prevalent cases that were diagnosed before the study began. It is not possible then to estimate relative risks consistently, unless the sampling fraction of cases and controls is known. But since odds ratios

$$\text{OR}(x) = \frac{O(x)}{O(0)} = \frac{\theta(x)/[1 - \theta(x)]}{\theta(0)/[1 - \theta(0)]} \quad (19)$$

can be estimated consistently also for the latter kind of studies (Cornfield 1951, Prentice and Pyke 1979, Skondal 2003) they may replace relative risks, both for AP due to marginal effects eq. (34) and interaction eq. (37).

Consider a subset  $J \subset \{1, \dots, p\}$  of factors, and a non-interaction model  $M$  among them. In Appendix A.3 we define a number of odds-ratio based quantities in terms of eq. (19) and

$$\text{OR}_{\text{rem}}(x; J, M) = \frac{\theta_{\text{rem}}(x; J, M)/[1 - \theta_{\text{rem}}(x; J, M)]}{\theta(0)/[1 - \theta(0)]}, \quad (20)$$

the odds ratio for exposure  $x$  when interaction among the factors in  $J$  has been removed, with  $\theta_{\text{rem}}(x; J, M)$  as in eq. (17). This includes the joint attributable proportion  $\text{AP}_{\text{OR}}(x; J)$  of the odds ratio for exposure  $x$  due to the factors in  $J$ , the corresponding joint attributable proportion  $\text{AP}_{\text{OR}}(J)$  when averaging the effect of all confounders, the attributable proportion  $\text{AP}_{\text{OR}}(x; J, M)$  of the odds ratio due to interaction among the factors in  $J$  for exposure  $x$ , and finally, the corresponding attributable proportion  $\text{AP}_{\text{OR}}(J, M)$  of the odds ratio due to interaction among the factors in  $J$ , when averaging over the confounding factors. When  $\text{AP}_{\text{OR}} > 0$ , it is the relative proportion of the odds ratio of disease attributed to joint effect or interaction among the variables in  $J$ . On the other hand, when  $\text{AP}_{\text{OR}} < 0$ , it is minus the proportion by which joint effect of interaction among the variables in  $J$  decreases the odds of developing the disease.

Odds ratios typically approximate relative risks well for rare diseases. When this is the case, the odds ratio based versions of the attributable proportion are close to the corresponding risk based quantities of Section 3. For the models studied by Kalilani and Atashili (2006), the odds ratio approximation was accurate for diseases with baseline risks up to 0.2. But sometimes the difference is notable, and depending on disease risks and mode of interaction the odds ratio based attributable proportion may be biased (Kalilani and Atashili 2006; Zou 2008).

The attributable proportion of the odds ratio due to interaction has a particularly explicit form for the additive odds model. For instance, when there are  $p=2$  risk factors and  $x=(1, 1)$ , we tacitly assume in the notation that  $J=\{1, 2\}$  (since otherwise  $\text{AP}_{\text{OR}}=0$ ) and find that

$$\text{AP}_{\text{OR}}(1, 1; \text{Additive odds}) = \frac{\text{OR}(1, 1) - [\text{OR}(0, 1) + \text{OR}(1, 0) - 1]_0^\infty}{\max\{\text{OR}(1, 1), [\text{OR}(0, 1) + \text{OR}(1, 0) - 1]_0^\infty\}} \quad (21)$$

is an odds ratio analogue of eqs (13)–(14). The corresponding traditional definition

$$\text{AP}_{\text{OR, trad}}(1, 1; \text{Additive odds}) = \frac{\text{OR}(1, 1) - [\text{OR}(0, 1) + \text{OR}(1, 0) - 1]}{\text{OR}(1, 1)}$$

can be found for instance in Rothman (2012).

## 5 Inference

We will outline a unified approach to estimate the attributable proportion eqs (4)–(5) or eqs (40)–(41), and the population attributable fraction [8] due to joint effects or interaction of an arbitrary subset  $J \subset \{1, \dots, p\}$  of factors. It is assumed that a saturated model  $M_{\text{sat}}$  is fit to data. For the risk-based quantities eqs (4)–(5) or eq. (8), the parameter vector  $\psi = (\psi_x)$  of  $M_{\text{sat}}$  has  $2^p$  components, one for each  $x$ . A linear parametrization of the risk function is

$$\theta(x) = \theta(x, \psi) = \psi_x. \quad (22)$$

Another possibility is to parametrize a logit transformation of risk. This was utilized by Hosmer and Lemeshow (1992) for the traditional definition [12] of AP of interaction with  $p=2$  risk factors and an additive model of no interaction. They noted that the disease probability for a subject with exposure  $x$  can be parametrized in two ways. Either

$$\theta(x) = \theta(x, \psi) = \frac{\exp\left(\sum_{y \leq x} \psi_y\right)}{1 + \exp\left(\sum_{y \leq x} \psi_y\right)}, \quad (23)$$



with  $y = (y_1, \dots, y_p)$  and  $y \leq x$  interpreted componentwise, i. e.  $y_i \leq x_i$  for  $i = 1, \dots, p$ , and  $\psi_x$  an interaction parameter defined for the  $|x| = \sum_i x_i$  nonzero risk factors of  $x$ . This is an ordinary logistic regression parametrization for a model with  $p$  predictors having 0 as their baseline levels, and interactions of all orders (Agresti 2013). If all nonzero exposure profiles  $x$  are viewed as different levels of one single predictor variable, then

$$\theta(x) = \theta(x, \psi) = \frac{\exp(\psi_0 + \psi_x)}{1 + \exp(\psi_0 + \psi_x)}, \quad (24)$$

with  $\psi_x$  the log odds of  $x$  compared to the baseline level 0. Since the two parametrizations eqs (23) and (24) only differ by a linear transformation, they give identical inference, although eq. (24) is often simpler. We can also express the odds ratio eq. (19) conveniently in terms of any of the two logistic parametrizations, either

$$\text{OR}(x) = \text{OR}(x, \psi) = \exp\left(\sum_{0 < y \leq x} \psi_y\right), \quad (25)$$

for eq. (23), or

$$\text{OR}(x) = \text{OR}(x, \psi) = \begin{cases} 1, & x = 0, \\ \exp(\psi_x), & x \neq 0, \end{cases} \quad (26)$$

for eq. (24). Neither eq. (25) nor eq. (26) involve the baseline parameter  $\psi_0$ , and therefore we redefine  $\psi$  to equal  $(\psi_x; x \neq 0)$  when quantities with odds ratios are estimated.

All quantities  $\text{AP}(x)$ ,  $\text{AP}_{\text{OR}}(x)$ ,  $\text{AP}$ ,  $\text{AP}_{\text{OR}}$  and  $\text{PAF}$  of interest are functions

$$\xi = \xi(\psi) = \frac{a(\psi) - b(\psi)}{\max[a(\psi), b(\psi)]}. \quad (27)$$

of the parameter vector  $\psi$ , with different choices of  $a$  and  $b$ , as summarized in Table 5. In order to estimate  $\xi$ , suppose we have data  $(x_i, Y_i)$ ,  $i = 1, \dots, N$  from  $N$  subjects, with  $\hat{\psi} = (\hat{\psi}_x)$  the accompanying maximum likelihood estimator of  $\hat{\psi}$ . If this is a cohort design or some other prospective study,  $\hat{\psi}$  maximizes a prospective likelihood, the probability of disease given exposure. For the linear parametrization (22) we have that

$$\hat{\psi}_x = \frac{N_{x1}}{N_x},$$

with  $N_{xy}$  the number of subjects  $i$  with  $x_i = x$  and  $Y_i = y$ , and  $N_x = N_{x0} + N_{x1}$ . For either of the two logit scale parametrizations (23)–(24), we use standard logistic regression software to compute  $\hat{\psi}$ . This is also possible for a case-control study, where sampling is on the response variable, and  $\hat{\psi}$  maximizes a retrospective likelihood, the probability of exposure given disease, as long as we focus on quantities that only involve odds ratios (Prentice and Pyke 1979). We estimate  $\xi$  by replacing  $\psi$  with  $\hat{\psi}$  in eq. (27), so that

$$\hat{\xi} = \xi(\hat{\psi}) = \frac{a(\hat{\psi}) - b(\hat{\psi})}{\max[a(\hat{\psi}), b(\hat{\psi})]}. \quad (28)$$

For large  $N$  one often approximates the distribution of  $\hat{\psi}$  by a multivariate normal  $N(\psi, \Sigma)$  with asymptotic covariance matrix  $\Sigma$ . The denominator of eq. (27) is not differentiable when  $a(\psi) = b(\psi)$ . But since the numerator is zero at such points,  $\xi(\psi)$  is still differentiable and  $\hat{\xi}$  approximately normally distributed  $N(\xi, \sigma^2)$  with asymptotic variance  $\sigma^2 = D\Sigma D^T$ ,  $D = \partial\xi(\psi)/\partial\psi$ , and  $D^T$  its transpose. As in Hosmer and Lemeshow (1992) we use the delta method to find a Wald type confidence interval

$$I = \left( \hat{\xi} - z_{1-\alpha/2} \hat{\sigma}, \hat{\xi} + z_{1-\alpha/2} \hat{\sigma} \right) \quad (29)$$

for  $\xi$  with asymptotic coverage probability  $1 - \alpha$ , where  $z_\beta$  is the  $\beta$ -quantile of a standard normal distribution, and

$$\hat{\sigma} = (\hat{D}\hat{\Sigma}\hat{D}^T)^{1/2} \quad (30)$$

the standard error of  $\hat{\xi}$ , with  $\hat{\Sigma}$  and  $\hat{D}$  estimates of  $\Sigma$  and  $D$ . For the linear parametrization (22),  $\hat{\Sigma}$  is a diagonal matrix with elements

$$\hat{\Sigma}_{xx} = \frac{\hat{y}_x(1 - \hat{y}_x)}{N_x}, \quad (31)$$

since observed risks are independent binomial proportions. For the logit parametrizations,  $\hat{\Sigma}$  is available from standard logistic regression software. A formula for  $\hat{D}$  is derived in Appendix B.

When  $\xi$  is close to one of its two boundary values  $-1$  and  $1$ , the distribution of  $\hat{\xi}$  will typically be skewed, unless the sample size  $N$  is very large. Since eq. (22) assumes a symmetric normal distribution of  $\hat{\xi}$ , the actual coverage probability  $\pi(1 - \alpha) = P(\xi \in I)$  may deviate substantially from the nominal  $1 - \alpha$ . This suggests that a conveniently chosen transformation of  $h(\hat{\xi})$  of  $\hat{\xi}$  will have a distribution closer to normal. When  $\xi \notin \{-1, 1\}$ , the logit type transformation  $h(x) = \log[(1 + x)/(1 - x)]$  gives a Wald confidence interval

$$I = \left( \frac{\exp(h(\hat{\xi}) - z_{1-\alpha/2}\tilde{\sigma}) - 1}{\exp(h(\hat{\xi}) - z_{1-\alpha/2}\tilde{\sigma}) + 1}, \frac{\exp(h(\hat{\xi}) + z_{1-\alpha/2}\tilde{\sigma}) - 1}{\exp(h(\hat{\xi}) + z_{1-\alpha/2}\tilde{\sigma}) + 1} \right), \quad (32)$$

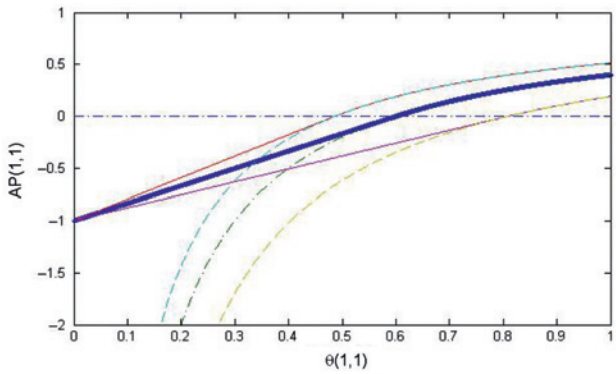
where  $\tilde{\sigma} = h'(\hat{\xi})\hat{\sigma} = 2\tilde{\sigma}/\{(1 + \hat{\xi})(1 - \hat{\xi})\}$ . Whereas the interval in eq. (29) may extend outside  $(-1, 1)$ , eq. (32) is always a subset of  $(-1, 1)$  as long as  $\hat{\xi} \in (-1, 1)$ . Other confidence interval approaches for  $\xi$  that deal with asymmetric distributions include the methods of variance estimates recovery (Zou 2008), likelihood ratio based intervals obtained from a saturated additive odds ratio model (Richardson and Kaufman 2009), nonparametric bootstrap (Assman et al. 1996), with accelerated bias correction (Zou 2008). In Appendix D we define a bias-corrected and accelerated version of the percentile bootstrap method.

The exposure distribution  $q(x) = P(X = x)$  is assumed to be known in eq. (28) for the population based AP, PAF and AP<sub>OR</sub>. In Appendix C we show how to compute standard error and confidence intervals when  $q$  is estimated from a prospective study.

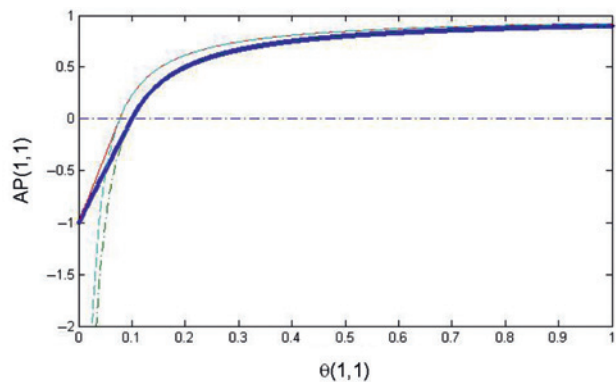
## 6 Numerical illustration and simulation

Figures 1–2 illustrate how normalizations eqs (2) and (4) affect AP(1, 1;  $M$ ) of interaction for two saturated models I and II with  $p = 2$  factors. Model I has two risk factors, whereas II has one risk and one preventive factor. As expected, the normalization has a large impact for both models when AP is negative. The model of no interaction influences the value of AP a lot for model I, and to some extent for model II. Table 2 gives values of the disease risk  $\theta_{\text{rem}}(x; J, M)$  when various types of interaction among the variables in  $J$  is removed. This is done for models I and II, as well as for a model III with  $p = 3$  co-factors.

Table 3 displays estimated coverage probabilities of confidence intervals for the odds ratio based attributable fraction AP<sub>OR</sub>( $x; J, M$ ) of interaction. Each coverage probability is estimated from a number of randomly drawn data sets with  $N/2$  cases and controls, generated from the two saturated models I and II with  $p = 2$  co-factors, and the model III with  $p = 3$ . We use two models  $M$  of no interaction, additive odds and multiplicative, and three types of confidence intervals; the delta method (29), the logit delta method (32) and the bootstrap method (48). It is seen that the accuracy of all three confidence intervals increase with  $N$ , and overall the bootstrap method has the best performance, although the simpler logit delta method works well for a wide range of models. Appendix F provides more details about the simulation study.



**Figure 1:** Plots of normalized attributable proportion  $AP(1, 1) = AP(1, 1; J, M)$  of interaction [37] (solid) and the corresponding traditional quantity [2] (dashed). The curves are functions of  $\theta(1, 1)$  for a saturated model with  $p = 2$  co-factors, whose other three risks  $\theta(0, 0) = 0.05$ ,  $\theta(1, 0) = 0.25$ ,  $\theta(0, 1) = 0.4$  are the same as for Model I of Table 2. Both variables are non-confounders ( $J = \{1, 2\}$ ), and three different no interaction models  $M$  are assumed; additive (thick line), additive odds (upper thin line) and multiplicative (lower thin line).



**Figure 2:** Plots of normalized attributable proportion  $AP(1, 1) = AP(1, 1; J, M)$  of interaction [37] (solid) and the corresponding traditional quantity [2] (dashed) versus  $\theta(1, 1)$  for a saturated model with  $p = 2$  risk factors, whose other three risks  $\theta(0, 0) = 0.1$ ,  $\theta(1, 0) = 0.05$ ,  $\theta(0, 1) = 0.15$  are the same as for Model II of Table 2. Both variables are non-confounders ( $J = \{1, 2\}$ ), and two different no interaction models  $M$  are assumed; additive (thick line), and multiplicative (thin line). The curves for the additive odds model (not shown) are very close to those of the additive model.

**Table 2:** Three saturated models I-III with  $p$  risk factors.

Sat. model	$p$					Additive	Add odds	Disj	Mult
		$\theta(0, 0)$	$\theta(1, 0)$	$\theta(0, 1)$	$\theta(1, 1)$	$\theta_{\text{rem}}(1, 1)$			
I	2	0.05	0.25	0.4	0.4	0.6	0.486	0.526	0.809
II	2	0.10	0.05	0.15	0.3	0.1	0.106	0.103	0.077
III	3	$p$	$\theta(0, 0, 0)$	$\theta(1, 0, 0)$	$\theta(0, 1, 0)$	$\theta(0, 0, 1)$	$\theta(1, 1, 0)$	$\theta(1, 0, 1)$	$\theta(0, 1, 1)$
									$\theta(1, 1, 1)$
			Additive	Add odds	Disj	Mult	Additive	Add odds	Disj
			$\theta_{\text{rem}}(1, 1, 0)$				$\theta_{\text{rem}}(1, 0, 1)$		
			0.400	0.362	0.378	0.491	0.250	0.270	0.261
			$\theta_{\text{rem}}(0, 1, 1)$				$\theta_{\text{rem}}(1, 1, 1)$		
			0.150	0.161	0.156	0.106	0.350	0.337	0.343

Note: For each model, the risks  $\theta(x)$  are displayed for all  $2^p$  exposure profiles  $x$ , and the predicted risks  $\theta_{\text{rem}}(x) = \theta_{\text{rem}}(x; J, M)$  for exposure profiles with at least two factors turned on. The latter are based on four different models  $M$  of no interaction among all co-factors ( $J = \{1, \dots, p\}$ ).

**Table 3:** Estimates  $\hat{\pi}(1-\alpha)$  of coverage probabilities for confidence intervals of odds ratio based APs of interaction.

Saturated model	AP <sub>OR</sub> ( $x; J, M$ )		Value	N	Delta		Logit delta		Bootstrap BC <sub>a</sub>	
	x	M			$\hat{\pi}(0.95)$	$\hat{\pi}(0.99)$	$\hat{\pi}(0.95)$	$\hat{\pi}(0.99)$	$\hat{\pi}(0.95)$	$\hat{\pi}(0.99)$
I	(1, 1)	Add odds	−0.296	500	0.906	0.959	0.932	0.980	0.956	0.991
				1,000	0.944	0.981	0.954	0.988	0.951	0.990
				10,000	0.950	0.990	0.950	0.990	0.948	0.989
I	(1, 1)	Mult	−0.842	500	0.917	0.956	0.954	0.993	0.957	0.993
				1,000	0.930	0.967	0.953	0.992	0.950	0.989
				10,000	0.949	0.988	0.951	0.990	0.951	0.989
II	(1, 1)	Add odds	0.725	500	0.938	0.975	0.978	0.998	0.952	0.989
				1,000	0.951	0.988	0.967	0.994	0.949	0.988
				10,000	0.950	0.990	0.952	0.991	0.950	0.989
II	(1, 1)	Mult	0.805	500	0.922	0.960	0.955	0.993	0.953	0.991
				1,000	0.934	0.972	0.952	0.991	0.950	0.989
				10,000	0.948	0.988	0.951	0.990	0.952	0.991
III	(1, 1, 0)	Add odds	0.149	1,000	0.901	0.960	0.923	0.977	0.957	0.992
				10,000	0.949	0.987	0.952	0.989	0.949	0.989
III	(1, 1, 0)	Mult	−0.309	1,000	0.876	0.934	0.922	0.975	0.958	0.991
				10,000	0.948	0.986	0.954	0.991	0.950	0.990
III	(1, 1, 1)	Add odds	0.943	1,000	0.916	0.957	0.958	0.993	0.959	0.993
				10,000	0.947	0.987	0.950	0.990	0.949	0.989
III	(1, 1, 1)	Mult	0.949	1,000	0.867	0.909	0.957	0.995	0.959	0.991
				10,000	0.939	0.976	0.950	0.991	0.947	0.990

Note: AP<sub>OR</sub>( $x; J, M$ ) in eq. (40) of Appendix A.3 quantifies interaction among all co-factors  $J = \{1, \dots, p\}$  for exposure  $x$ , with data generated from the three saturated models of Table 2, with Additive odds or Multiplicative models of no interaction. The confidence intervals with nominal coverage  $1-\alpha$  use the delta method (29), the logit delta method (32) or the bias corrected and accelerated bootstrap method (48) with  $B=1000$  resamples. The delta methods (bootstrap) employ 1,00,000 (10,000) retrospectively simulated data sets from the saturated model, with  $N/2$  cases and controls.

## 7 A real data set

MS is a complex and inflammatory disease causing damage to the central nervous system. A number of family studies indicate a genetic component of the disease, with a major contribution from variants at the human leukocyte antigen (HLA) complex. It is well known that presence of allele 15 of the HLA-DRB1 gene is a risk factor, whereas allele 02 of the HLA-A gene has a protective effect (Cree 2014). Several data sets reveal that environmental factors, in particular smoking, impact the aetiology of the disease as well. A recent paper of Hedström et al. (2011) demonstrated pairwise positive additive interaction between these three factors for a Scandinavian case-control data set with 843 cases and 1209 controls. Here we will redo parts of the analysis, using our novel measures AP<sub>OR</sub> of joint effects and interaction. As in Hedström et al. (2011), we use a binary coding  $x_{DRB1*15} = 1$  for an individual with genotype 15/X or 15/15,  $x_{A*02} = 1$  if the genotype is Z/Z, and  $x_{smoking} = 1$  for a current smoker, with  $X$  and  $Z$  referring to any allele different from HLA-DRB1\*15 and HLA-A\*02 respectively. For simplicity, we will not adjust for the confounding effect of age, gender, residential area and ancestry, as in Hedström et al. (2011). Table 4 summarizes results for different models, three of which have one single factor, one with both genetic factors included, and one model with all three factors present.

Each one-factor model quantifies the marginal effects of a single covariate, without taking the confounding effect of the other two into account. DRB1\*15, A\*02 and smoking are risk factors, since the estimated odds ratios  $\widehat{OR}(1)$  between MS and each factor equals 3.47, 1.63 and 1.58 when the other two

**Table 4:** Analysis of the Multiple Sclerosis data set of Hedström et al. (2011). This includes point estimates and confidence intervals of the odds ratio based attributable proportion  $AP_{OR}$  for joint effects and interaction, and point estimates of the odds ratio  $OR$  before and odds ratio  $OR_{rem}$  after the joint effect or interaction has been removed.

Factors	$x$	$J$	$M$	$\widehat{AP}_{OR}$	$\widehat{OR}$	$\widehat{OR}_{rem}$	$I_{Ld}$	$I_{BC_a}$
DRB1*15	1	{1}		0.711	3.465	1.000	(0.654,0.761)	(0.652,0.760)
A*02	1	{1}		0.387	1.630	1.000	(0.273,0.490)	(0.267,0.486)
smoking	1	{1}		0.366	1.578	1.000	(0.232,0.487)	(0.224,0.479)
(DRB1*15, A*02)	(1, 0)	{1}		0.718	3.542	1.000	(0.633,0.785)	(0.630,0.784)
	(0, 1)	{2}		0.396	1.653	1.000	(0.234,0.535)	(0.222,0.530)
	(1, 1)	{1}		0.704	5.576	1.653	(0.618,0.772)	(0.614,0.771)
	(1, 1)	{2}		0.365	5.576	3.542	(0.178,0.526)	(0.158,0.516)
	(1, 1)	{1, 2}		0.821	5.576	1.000	(0.767,0.863)	(0.766,0.863)
	(1, 1)	{1, 2}	Add odds	0.248	5.576	4.196	(0.038,0.436)	(0.015,0.428)
	(1, 1)	{1, 2}	Mult	-0.048	5.576	5.856	(-0.383,0.299)	(-0.347,0.273)
	(1, 0, 0)	{1}		0.728	3.678	1.000	(0.632,0.802)	(0.628,0.801)
	(0, 1, 0)	{2}		0.351	1.542	1.000	(0.149,0.525)	(0.131,0.517)
(DRB1*15, A*02, smoking)	(1, 1, 0)	{1}		0.678	4.782	1.542	(0.570,0.763)	(0.565,0.760)
	(1, 1, 0)	{2}		0.231	4.782	3.678	(-0.023,0.457)	(-0.065,0.436)
	(1, 1, 0)	{1, 2}		0.791	4.782	1.000	(0.718,0.847)	(0.717,0.846)
	(1, 1, 0)	{1, 2}	Add odds	0.118	4.782	4.220	(-0.158,0.376)	(-0.169,0.362)
	(1, 1, 0)	{1, 2}	Mult	-0.157	4.782	5.670	(-0.486,0.212)	(-0.457,0.232)
	(1, 1, 1)	{1}		0.803	13.340	2.635	(0.659,0.891)	(0.648,0.891)
	(1, 1, 1)	{2}		0.691	13.340	4.148	(0.457,0.835)	(0.429,0.833)
	(1, 1, 1)	{1, 2}		0.900	13.340	1.335	(0.822,0.945)	(0.817,0.946)
	(1, 1, 1)	{1, 2}	Add odds	0.593	13.340	5.447	(0.309,0.780)	(0.271,0.781)
	(1, 1, 1)	{1, 2}	Mult	0.389	13.340	8.183	(-0.146,0.748)	(-0.255,0.724)
	(1, 1, 1)	{1, 2, 3}		0.925	13.340	1.000	(0.876,0.956)	(0.875,0.956)
	(1, 1, 1)	{1, 2, 3}	Add odds	0.660	13.340	4.555	(0.442,0.804)	(0.425,0.804)
	(1, 1, 1)	{1, 2, 3}	Mult	0.435	13.340	7.572	(-0.084,0.768)	(-0.222,0.741)

Note: The two genetic variables DRB1\*15 and A\*02 encode genotypes of human leukocyte antigens in a binary way, as described in the text, and smoking is an indicator of current smoking. All three variables are binary risk factors, with level 1 indicating increased risk. Column 1 lists the  $p \in \{1, 2, 3\}$  co-factors that are included in the model, column 2 displays for each factor whether risk is present (1) or absent (0), and column 3 shows which co-factors ( $J \in \{1, \dots, p\}$ ) that are included in  $AP_{OR}$ , see eq. (40) for a definition. The fourth column displays whether joint effects (there is no model  $M$  of no interaction) or interaction ( $M$  is specified) between the factors in  $J$  is estimated. The fifth column either shows point estimates of  $AP_{OR}(x; J)$  in eq. (40) for joint effects, or of  $AP_{OR}(x; J, M)$  in eq. (42) for interaction. The sixth column contains point estimates of  $OR(x)$  in eq. (19), and the seventh column either lists estimates of  $OR(x_j)$  for joint effects, or of  $OR_{rem}(x; J, M)$  in eq. (20) for interaction. The eighth column is the Logit delta based confidence interval in eq. (32), and the last column the bootstrapped  $BC_a$  interval of eq. (48), based on  $B = 20000$  replicates. Both confidence intervals have a nominal coverage probability of  $1 - \alpha = 0.95$ .

factors are not accounted for, see also Table 1 of Hedström et al. (2011). The fraction of odds ratio attributable to each factor is

$$\widehat{AP}_{OR}(1; \{1\}) = \frac{\widehat{OR}(1) - 1}{\max[\widehat{OR}(1), 1]} = \begin{cases} 0.711, & \text{for DRB1*15,} \\ 0.387, & \text{for A*02,} \\ 0.366, & \text{for smoking,} \end{cases}$$

and the accompanying confidence intervals are displayed in the first three rows of Table 4. This does not imply that more than 70 % of the odds ratio between MS and various risk factors is attributable to DRB1\*15, only that this is the case when no other co-factors or sufficient causes than DRB1\*15 are taken into account, cf. Lekman et al. (2014) for a more extensive discussion. We also notice from Table 4 that the estimated  $AP_{OR}$  for DRB1\*15 changes very little, to 0.718 and 0.704, when the confounding effect of A\*02 is controlled

for at levels 0 and 1, and to 0.728, 0.678 and 0.803 when A\*O2 and smoking are controlled for jointly, at levels (0,0), (1,0) and (1,1).

Next we estimate the joint effect of two or more variables, and notice from Table 4 that the estimated proportion of odds ratio attributable to DRB1\*15 and A\*O2, is 0.821, 0.791 and 0.900 when smoking is not accounted for, among non-smokers and among smokers respectively. The estimated  $AP_{OR}$  is even larger, 0.925, for all three variables jointly. This corresponds to an odds ratio that drops from 13.340 to 1 when the joint effect of all factors is removed.

The framed numbers of Table 4 are estimates of additive  $AP_{OR}$  of interaction between the two genetic factors DRB1\*15 and A\*O2 under various smoking scenarios. It is seen that  $AP_{OR}$  varies substantially; 0.248, 0.118 and 0.593, depending on whether smoking is not accounted for, if only non-smokers are considered, or if only smokers are included. In the first case the odds ratio drops from 5.576 to 4.196 when interaction between the two factors is *removed*, in the second case it drops from 4.782 to 4.220, and in the third case from 13.340 to 5.447. The corresponding confidence intervals in Table 4 agree fairly well with those in Table 5 of Hedström et al. (2011). We are also able to estimate the proportion of additive interaction (0.660) among all three factors jointly, which corresponds to an odds ratio that decreases from 13.340 to 4.555 when interaction among the three variables is removed. The corresponding fractions of odds ratio attributable to multiplicative interaction between two or three factors are given in Table 4 as well, with accompanying confidence intervals. For instance, the proportion of multiplicative interaction between DRB1\*15 and A\*O2 is  $-0.048$  when smoking is not accounted for, corresponding to an odds ratio that decreases from 5.856 to 5.576 when interaction among the two variables is *added*. The multiplicative interaction is often much weaker than the additive, in particular for the two genetic risk factors. This implies that the joint effects estimates of  $AP_{OR}$  are less sensitive to the levels of the confounding variables, see formula [51] of Appendix E.

Suppose, finally, that the two labels of A\*O2 are switched (before looking at data). Then this factor becomes protective, with  $x_{A*O2} = 1$  if the genotype is 2/Z or 2/2, and the estimated odds ratio between A\*O2 and MS changes from 1.63 to  $1/1.63$ . Consequently, the estimated marginal attributable proportion

$$\widehat{AP}_{OR}(1; \{1\}) = \frac{1/1.63 - 1}{\max(1/1.63, 1)} = -0.386$$

for A\*O2 becomes negative, with a 95 % confidence interval  $(-0.490, -0.273)$  for the logit delta method, and  $(-0.487, -0.267)$  for the bootstrap method. The proportion of odds ratio due to additive interaction between DRB1\*15 and A\*O2 becomes negative as well  $(-0.280)$  when smoking is not controlled for, with 95 % confidence intervals  $(-0.482, -0.051)$  and  $(-0.468, -0.022)$  for the logit delta and bootstrap methods.

**Table 5:** Examples of quantities  $\xi$  in eq. (27).

$\xi$	$a(\psi)$	$b(\psi)$	Type of quantity
$AP(x; J)$	$\theta(x, \psi)$	$\theta(x_J, \psi)$	Joint effect
$AP(J)$	$\sum_x \theta(x', \psi) q(x)$	$\sum_x \theta(x_J, \psi) q(x)$	
$PAF(J)$	$\sum_x \theta(x, \psi) q(x)$	$\sum_x \theta(x_J, \psi) q(x)$	
$AP_{OR}(x; J)$	$OR(x, \psi)$	$OR(x_J, \psi)$	
$AP_{OR}(J)$	$\sum_x OR(x', \psi) q(x)$	$\sum_x OR(x_J, \psi) q(x)$	
$AP(x; J, M)$	$\theta(x, \psi)$	$\theta_{rem}(x, \psi; J, M)$	Interaction
$AP(J, M)$	$\sum_x \theta(x', \psi) q(x)$	$\sum_x \theta_{rem}(x', \psi; J, M) q(x)$	
$PAF(J, M)$	$\sum_x \theta(x, \psi) q(x)$	$\sum_x \theta_{rem}(x, \psi; J, M) q(x)$	
$AP_{OR}(x; J, M)$	$OR(x, \psi)$	$OR_{rem}(x, \psi; J, M)$	
$AP_{OR}(J, M)$	$\sum_x OR(x', \psi) q(x)$	$\sum_x OR_{rem}(x', \psi; J, M) q(x)$	

Note: All risk based quantities involve the  $\theta(x, \psi)$ , given by either of eqs (22)–(24). The odds ratio based quantities involve  $OR(x, \psi)$  in eqs (25) or (26). The quantity  $\theta_{rem}(x, \psi; J, M)$  in rows 6–8 refers to the risk function [17] when interaction among the variables in  $J$  has been removed according to model  $M$ , see Table 6. The quantity  $OR_{rem}(x, \psi; J, M)$  in the last two rows refers to the odds ratio [20] when interaction among the variables in  $J$  has been removed, see Table 7. The exposure distribution in the study population is  $q(x)$ .



## 8 Discussion

We studied the attributable proportion (AP) and its population counterpart (PAF) for models of disease that involve a number of binary predictor variables. The definition of AP and PAF was extended in several ways. First, we put AP and PAF for joint effects, marginal effects and interaction into a general framework. Second, we introduced a novel normalization of AP and PAF between  $-1$  and  $1$ , with a natural interpretation in terms of proportion of risk attributable to removing or adding joint or interaction effects. Third, we allowed for an arbitrary number of co-factors and subsets thereof for which exposure is of main concern, whereas the other confounders are controlled for. Fourth, our framework incorporates not only additive models of no interaction, but also various other models. Fifth, we outlined a general approach for point estimates and various types of confidence intervals for prospective or case-control studies.

Several extensions are possible. First, we assume that a null model  $M$  of no interaction is embedded into a larger saturated model  $M_{\text{sat}}$ . For a logit parametrization of risk, the multiplicative model is a linear subspace of  $M_{\text{sat}}$ , but other null models are more complicated curved exponential families (Lehmann and Casella 1998). An advantage of our approach is that no parameter estimates are needed for  $M$ , whereas those for  $M_{\text{sat}}$  are straightforward; binomial proportions or output from standard logistic regression software. Apart from this we only need formulas for how  $M$  predicts risks or odds, and the derivatives of these predictions with respect to model parameters. Our approach requires data for all  $2^p$  exposure profiles in order to estimate the parameters of the saturated model. If such data is not available, it is better to use a smaller alternative model than  $M_{\text{sat}} \setminus M$ , or to reduce  $p$ .

Second, our confidence intervals for the attributable proportion can be used to define Wald type goodness of fit tests, which reject no joint effects or interaction among the tested factors whenever the confidence intervals do not include  $0$ . One may also employ a likelihood ratio test between the smaller null model and  $M_{\text{sat}} \setminus M$ . But this approach is less feasible when the null model is a complex no interaction model  $M$ , with a maximum likelihood estimator that is possibly much more complicated than the estimate for  $M_{\text{sat}}$ .

Third, there is sometimes no *a priori* knowledge of which co-factor levels to use as baseline. For instance, among a pair of SNPs it may not be known which one that increases disease risk. Then one may use the symmetrized version

$$\text{AP}^{\text{symm}} = \frac{|\theta(1,1) + \theta(0,0) - [\theta(0,1) + \theta(1,0)]|}{\max[\theta(1,1) + \theta(0,0), \theta(0,1) + \theta(1,0)]} \quad (33)$$

to quantify departures from an additive model of no interaction, since it is invariant with respect to labeling of the two factors, and takes a value between  $0$  and  $1$ . A value of  $0$  indicates no additive interaction, and positive values that additive interaction is present. Since the labeling of at least one risk factor is arbitrary, we cannot decide whether interaction departs negatively or positively from the additive model. The odds ratio version  $\text{AP}_{\text{OR}}^{\text{symm}}$  replaces  $\theta(x_1, x_2)$  by  $\text{OR}(x_1, x_2)$  in eq. (33). Because of the singularity at  $0$ , it is not straightforward to produce confidence intervals for the risk or odds ratio based versions of  $\text{AP}^{\text{symm}}$  though, and a Bayesian approach might be preferable.

Fourth, attributable fractions have been used for models where some or all of the  $p$  covariates are continuous, see Eide (2001) and references therein. Our new definitions eqs. (4), (5) and (8) of the attributable proportion and population attributable fraction apply to these settings as well. However, the parameter vector  $\psi$  becomes infinite dimensional, and a challenging topic for future research is to modify the inference procedure of Section 5 to this setting. See also Moore (2004) and Moore et al. (2006) for nonparametric machine learning and data mining approaches to quantify interaction for binary predictors.

Fifth, the synergy index is a measure of additive interaction of two risk factors  $x$ . It has been advocated by Skrdal (2003) as less sensitive to additional covariates associated to the disease but not to  $x = (x_1, x_2)$ , see also Greenland (1993). A generalized version

$$S(x; J, M) = \frac{\theta(x) - \theta(x_J)}{\theta_{\text{rem}}(x; J, M) - \theta(x_J)} = \frac{\text{RR}(x) - \text{RR}(x_J)}{\text{RR}_{\text{rem}}(x; J, M) - \text{RR}(x_J)}$$

of the synergy index could be defined for any no interaction model  $M$  and number of risk factors, or a version  $S_{OR}(x; J, M)$  that uses odds ratios instead of relative risks. They predict increased risk or OR, with values exceeding 1, when interaction among the variables in  $J$  is positive or synergetic. Point estimates and confidence intervals for  $S(x; J, M)$  and  $S_{OR}(x; J, M)$  are obtained as in Section 5. The latter assume asymptotic normality of  $\hat{S}$ , or of  $\log(\hat{S})$ , if it is known *a priori* that  $S > 0$  (Rothman 1976b).

In summary, we argue our unified theory for AP and PAF of joint effects or interactions among an arbitrary number of factors contributes with significant knowledge to understand synergies of biological and environmental factors that influence health and disease.

## Appendix A. General definitions of attributable proportions and its generalizations in terms of risk ratios and odds ratios

### A.1 Marginal and joint effects

Consider a subset  $J \subset \{1, \dots, p\}$  of factors, and an exposure vector  $x = (x_1, \dots, x_p)$ . The joint attributable proportion of risk in eq. (4) due to the factors in  $J$ , can be expressed in terms of the risk ratio [16], as

$$AP(x; J) = \frac{RR(x) - RR(x_J)}{\max[RR(x), RR(x_J)]}, \quad (34)$$

where  $x_J$  is the vector for which exposure to all factors in  $J$  has been removed (see Section 3.2). This follows by first substituting  $\theta_{rem}(x) = \theta(x_J)$  into eq. (4), and then dividing the numerator and denominator of this equation by  $\theta(0)$ .

In order to define the joint attributable proportion of risk eq. [5] among the factors in  $J$ , we associate to each  $x$  another vector  $x^J = (x_1^J, \dots, x_p^J)$ , whose exposure to all factors  $i \in J$  has been *added* or turned on ( $x_i^J = 1$ ), whereas all confounding factors  $i \notin J$  remain at the levels specified by  $x$  ( $x_i^J = x_i$ ). Let also  $X$  refer to the exposure vector of a randomly chosen individual. Then eq. (5) takes the form

$$AP(J) = \frac{E[RR(X^J)] - E[RR(X_J)]}{\max\{E[RR(X^J)], E[RR(X_J)]\}}, \quad (35)$$

with expectation referring to the average effect of the confounders. This follows by substituting  $\theta_{add}(x) = \theta(x^J)$  into eq. (5), and then dividing the numerator and denominator of this equation by  $\theta(0)$ . The joint population attributable fraction [8] due to the factors in  $J$ , is similarly expressed as

$$PAF(J) = \frac{E[RR(X)] - E[RR(X_J)]}{\max\{E[RR(X)], E[RR(X_J)]\}}. \quad (36)$$

### A.2 Interaction

The versions eq. (4) and (5) of the attributable fraction due to interaction among the variables in  $J$  can be expressed in terms of the risk ratios [16] and [18], as

$$AP(x; J, M) = \frac{RR(x) - RR_{rem}(x; J, M)}{\max[RR(x), RR_{rem}(x; J, M)]} \quad (37)$$

for a fixed exposure vector  $x$ , and as

$$AP(J; M) = \frac{E[RR(X^J)] - E[RR_{rem}(X^J; J, M)]}{\max\{E[RR(X^J)], E[RR_{rem}(X^J; J, M)]\}} \quad (38)$$

when averaging over confounders. The population attributable fraction [8] due to interaction among the factors in  $J$  is similarly rewritten as

$$\text{PAF}(J, M) = \frac{E[\text{RR}(X)] - E[\text{RR}_{\text{rem}}(x; J, M)]}{\max\{E[\text{RR}(X)], E[\text{RR}_{\text{rem}}(x; J, M)]\}}. \quad (39)$$

### A.3 Odds ratio based quantities

The odds ratio versions of attributable proportion use odds ratios [19] and [20] instead of relative risks in eqs (34)–(35) and eqs (37)–(38). By making such a substitution, we find that

$$\text{AP}_{\text{OR}}(x; J) = \frac{\text{OR}(x) - \text{OR}(x_J)}{\max[\text{OR}(x), \text{OR}(x_J)]}, \quad (40)$$

$$\text{AP}_{\text{OR}}(J) = \frac{E[\text{OR}(X^J)] - E[\text{OR}(X_J)]}{\max\{E[\text{OR}(X^J)], E[\text{OR}(X_J)]\}}, \quad (41)$$

$$\text{AP}_{\text{OR}}(x; J, M) = \frac{\text{OR}(x) - \text{OR}_{\text{rem}}(x; J, M)}{\max[\text{OR}(x), \text{OR}_{\text{rem}}(x; J, M)]} \quad (42)$$

and

$$\text{AP}_{\text{OR}}(J; M) = \frac{E[\text{OR}(X^J)] - E[\text{OR}_{\text{rem}}(X^J; J, M)]}{\max\{E[\text{OR}(X^J)], E[\text{OR}_{\text{rem}}(X^J; J, M)]\}}. \quad (43)$$

## Appendix B. Delta based confidence interval

In this appendix we describe how to produce delta based confidence intervals for  $\xi$  in eq. (27). This quantity either corresponds to one of the five notions  $\text{AP}(x; J)$ ,  $\text{AP}(J)$ ,  $\text{PAF}(J)$ ,  $\text{AP}_{\text{OR}}(x; J)$ ,  $\text{AP}_{\text{OR}}(J)$  of attributable proportion or fraction for joint effects in eqs (34)–(36) and eqs (40)–(41), and the corresponding quantities  $\text{AP}(x; J, M)$ ,  $\text{AP}(J, M)$ ,  $\text{PAF}(J, M)$ ,  $\text{AP}_{\text{OR}}(x; J, M)$ ,  $\text{AP}_{\text{OR}}(J, M)$  for interaction in eqs (37)–(39) and eqs (42)–(43). For each of these quantities,  $\xi = \xi(\psi)$  involves the parameter vector  $\psi$  and the two functions  $a = a(\psi)$  and  $b = b(\psi)$ , as shown in Table 5. The population based quantities  $\text{AP}(J)$ ,  $\text{PAF}(J)$ ,  $\text{AP}_{\text{OR}}(J)$  of joint effects, and  $\text{AP}(J; M)$ ,  $\text{PAF}(J; M)$ ,  $\text{AP}_{\text{OR}}(J; M)$  of interaction, involve the exposure distribution  $q$  as well. In this appendix it is assumed to be known, but in Appendix C we extend the analysis to situations where  $q$  is estimated from a prospective study.

Let  $\hat{\psi}$  the the maximum likelihood estimator of  $\psi$ , and  $\hat{\xi} = \xi(\hat{\psi})$  the accompanying plug in estimate of  $\xi$ . The two delta based confidence intervals (29) and (32) involve the standard error eq. (30) of  $\hat{\xi}$ . In order to find the standard error we need an estimate  $\hat{\Sigma}$  of the covariance matrix of  $\hat{\psi}$ . For a linear parametrization of the risk function, eq. (31) describes how to compute  $\hat{\Sigma}$ , and for a logit parametrization, it is available from standard logistic regression software. The other quantity  $\hat{D}$  of the standard error is an estimate of  $D = \partial \xi(\psi) / \partial \psi$ , the derivative of  $\xi$  with respect to the parameter vector  $\psi$ . In order to compute  $D$  we introduce  $A = da(\psi) / d\psi$  and  $B = db(\psi) / d\psi$ . Differentiating eq. (27) with respect to  $\psi$ , we get

$$\begin{aligned} D(\psi) &= \frac{A(\psi) - B(\psi)}{\max[a(\psi), b(\psi)]} \\ &\quad - 1_{\{a(\psi) > b(\psi)\}} \frac{[a(\psi) - b(\psi)]A(\psi)}{a(\psi)^2} \\ &\quad - 1_{\{a(\psi) < b(\psi)\}} \frac{[a(\psi) - b(\psi)]B(\psi)}{b(\psi)^2}, \end{aligned} \quad (44)$$

with  $1_\Omega$  the indicator function of  $\Omega$ , which equals 1 when  $\Omega$  holds and 0 otherwise. Although there is a point of discontinuity of the derivative of the denominator of eq. (27),  $D$  is still differentiable, since the numerator vanishes at such points. We estimate eq. (44) by

$$\hat{D} = D(\hat{\psi}),$$

and insert this estimate into eq. (30) in order to obtain the standard error of  $\hat{\xi}$ .

The two parametrizations  $\psi_1 = (\psi_x^1; x \in \{0, 1\}^p)$  and  $\psi_2 = (\psi_x^2; x \in \{0, 1\}^p)$  of the saturated logistic regression model in eqs (23) and (24), only differ by a linear transformation

$$\psi_x^2 = \begin{cases} \psi_0^1, & x = 0, \\ \sum_{0 < y \leq x} \psi_y^1, & x \neq 0, \end{cases}$$

or equivalently

$$\psi_2 = \psi_1 C, \quad (45)$$

for a square matrix  $C$  of order  $2^p$ . This implies  $\xi_1(\psi_1) = \xi_2(\psi_1 C)$ ,  $\hat{\psi}_2 = \hat{\psi}_1 C$ ,  $\hat{\Sigma}_2 = C^T \hat{\Sigma}_1 C$  and  $\hat{D}_1 = \hat{D}_2 C^T$ . Therefore the two parametrizations give identical point estimates  $\hat{\xi} = \xi_i(\hat{\psi}_i)$ , standard errors  $\hat{\sigma} = (\hat{D}_i \hat{\Sigma}_i \hat{D}_i^T)^{1/2}$ , delta based and logit delta based confidence intervals. It can be shown, by a Taylor expansion argument, that the linear parametrization (22) gives identical standard errors and (logit) delta based confidence intervals as well.

**Example 1 (Odds ratio based AP of interaction between two predictors.)** We illustrate the two logit parametrizations for  $p = 2$  factors, where  $\xi = \text{AP}_{\text{OR}}[(1, 1); J, M]$  quantifies interaction between two predictors  $J = \{1, 2\}$  by means of a no interaction model  $M$ . We suppress  $J$  in the notation and notice from Table 5 that eq. (27) holds with  $a(\psi) = \text{OR}[(1, 1), \psi]$  and  $b(\psi) = \text{OR}_{\text{rem}}[(1, 1), \psi; M]$ . Since none of the two logit parametrizations include the baseline risk  $\psi_{00}$ , we write their parameter vectors as  $\psi_i = (\psi_{01}^i, \psi_{10}^i, \psi_{11}^i)$  for  $i = 1, 2$ . It follows from eqs (25)–(26) that these reduced parameter vectors satisfy the same type of linear relationship [45] as the non-reduced vectors, with

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

so that  $\psi_{01}^1 = \psi_{01}^2 = \psi_{01}$ ,  $\psi_{10}^1 = \psi_{10}^2 = \psi_{10}$ , and  $\psi_{11}^2 = \psi_{01} + \psi_{10} + \psi_{11}^1$ . From eqs (25)–(26) we find that

$$a(\psi_1) = \exp(\psi_{01}^1 + \psi_{10}^1 + \psi_{11}^1),$$

$$a(\psi_2) = \exp(\psi_{11}^2),$$

are functions of the chosen parametrization, but not of the no interaction model  $M$ . Taking the derivative with respect to  $\psi$ , it follows that

$$A(\psi_1) = a(\psi_1)(1, 1, 1),$$

$$A(\psi_2) = a(\psi_2)(0, 0, 1).$$

By Table 7,  $b(\psi)$  only depends on first two coordinates  $\psi_{01}$  and  $\psi_{10}$  of the parameter vector, and hence is the same for both parametrizations. We have that

$$b(\psi_1) = b(\psi_2) = \begin{cases} \exp(\psi_{01}) + \exp(\psi_{10}) - 1, & M = \text{Additive odds}, \\ \exp(\psi_{01} + \psi_{10}), & M = \text{Multiplicative}. \end{cases}$$

Differentiating with respect to  $\psi$  we obtain

$$B(\psi_1) = B(\psi_2) = \begin{cases} [\exp(\psi_{01}), \exp(\psi_{10}), 0], & M = \text{Additive odds,} \\ b(\psi_2)(1, 1, 0), & M = \text{Multiplicative.} \end{cases}$$

## Appendix C. Delta based confidence interval with unknown exposure distribution

In the definition of  $\hat{\xi}$  for the population based quantities AP, PAF and  $AP_{OR}$ , it is assumed that the regression parameters  $\psi$  are unknown, whereas the exposure distribution  $q$  is known. Since this is rarely the case,  $q$  has to be estimated. Fortunately it is possible to enlarge the parameter vector  $\psi$  to include  $q$  as well. We will outline how the standard error  $\hat{\sigma}$  for a prospective study can be computed, so that formulas [29] and [32] still apply for the two delta based confidence intervals.

Suppose the exposure distribution  $q = \{q(x); x \in \{0, 1\}^p\}$  is unknown, represented as a probability vector with  $2^p - 1$  free parameters. This gives an enlarged parameters vector

$$\bar{\psi} = (\psi, q)$$

that includes effect as well as exposure distribution quantities. Since AP, PAF and  $AP_{OR}$  all depend on the exposure distribution, we write these quantities as

$$\xi = \xi(\bar{\psi}) = \frac{a(\bar{\psi}) - b(\bar{\psi})}{\max[a(\bar{\psi}), b(\bar{\psi})]} \quad (46)$$

instead of eq. (27). For a prospective study with data  $\{(x_i, Y_i); i = 1, \dots, N\}$ , we estimate the exposure distribution as

$$\hat{q}(x) = \frac{N_x}{N},$$

for all binary vectors  $x$  of length  $p$ , with  $N_x$  the number of subjects  $i$  with  $x_i = x$ . Let  $\hat{\bar{\psi}} = (\hat{\psi}, \hat{q})$  be the vector of estimated effect and exposure distribution parameters, with  $\hat{q} = \{\hat{q}(x); x \in \{0, 1\}^p\}$ . Insertion into eq. (46) gives a plug in estimate

$$\hat{\xi} = \xi(\hat{\bar{\psi}}) = \frac{a(\hat{\bar{\psi}}) - b(\hat{\bar{\psi}})}{\max[a(\hat{\bar{\psi}}), b(\hat{\bar{\psi}})]}$$

of  $\xi$ . The two delta based confidence intervals eqs (29) and (32) still apply, with a standard error computed as

$$\hat{\sigma} = \left( \hat{\bar{D}} \hat{\bar{\Sigma}} \hat{\bar{D}}^T \right)^{1/2} \quad (47)$$

instead of eq. (30). Here  $\hat{\bar{D}} = \bar{D}(\hat{\bar{\psi}})$  is an estimate of  $\bar{D}(\bar{\psi}) = d\xi(\bar{\psi})/d\bar{\psi}$ , computed similarly as in eq. (44), and

$$\hat{\bar{\Sigma}} = \begin{pmatrix} \hat{\Sigma} & 0 \\ 0 & \hat{\Sigma}_q \end{pmatrix}$$

is an estimate of the covariance matrix  $\bar{\Sigma}$  of the enlarged parameter vector, using the fact that the effect and exposure parameters are asymptotically orthogonal. Since  $N\hat{q}$  has a multinomial distribution, the estimated covariance matrix of the exposure parameters  $\hat{q}$ , has elements

$$\hat{\Sigma}_{q, xy} = \begin{cases} \hat{q}(x)[1 - \hat{q}(x)]/N, & x = y, \\ -\hat{q}(x)\hat{q}(y)/N, & x \neq y. \end{cases}$$

## Appendix D. Bootstrap $BC_\alpha$ confidence interval

We use nonparametric bootstrap (Efron and Tibshirani 1993) and define for each resampled data set  $(x_1^*, Y_1^*), \dots, (x_N^*, Y_N^*)$  a corresponding estimate  $\hat{\xi}^* = \xi(\hat{\psi}^*)$ , where

$$\hat{\psi}^* = \hat{\psi}(\{x_i^*, Y_i^*\}).$$

The resampled data points  $(x_i^*, Y_i^*)$  are drawn randomly with replacement from the original data set  $\{(x_i, Y_i)\}_{i=1}^N$  for a prospective study, and separately with replacement within controls and cases for a retrospective study, keeping the total number of controls and cases fixed in each resample. The bias-corrected accelerated percentile method of Efron (1987) is based on  $B$  independent replicates  $\hat{\xi}_1^*, \dots, \hat{\xi}_B^*$  of  $\hat{\xi}^*$ , and a confidence interval

$$I = \left( \hat{\xi}_{([B\pi_1])}^*, \hat{\xi}_{([B\pi_2])}^* \right) \quad (48)$$

with asymptotic coverage  $1 - \alpha$ , where  $\hat{\xi}_{(1)}^* \leq \dots \leq \hat{\xi}_{(B)}^*$  are the ordered  $\hat{\xi}_i^*$ ,

$$\pi_1 = \Phi \left( \hat{z} + \frac{\hat{z} + Z_{\alpha/2}}{1 - \hat{a}(\hat{z} + Z_{\alpha/2})} \right),$$

$$\pi_2 = \Phi \left( \hat{z} + \frac{\hat{z} + Z_{1-\alpha/2}}{1 - \hat{a}(\hat{z} + Z_{1-\alpha/2})} \right),$$

$\Phi$  is the cumulative distribution function of a standard normal,

$$\hat{z} = Z \left\{ \sum_{i=1}^B 1_{\hat{\xi}_i^* \leq \hat{\xi}} / B \right\}$$

a bias correction term, and

$$\hat{a} = \frac{\sum_{i=1}^N (\hat{\xi}_{(\cdot)} - \hat{\xi}_{(-i)})^3}{6 \left\{ \sum_{i=1}^N (\hat{\xi}_{(\cdot)} - \hat{\xi}_{(-i)})^2 \right\}^{3/2}},$$

an acceleration term that is a rescaled estimate of the skewness of the influence function of  $\hat{\xi}$  (Hampel et al. 1986), with  $\hat{\xi}_{(-i)}$  the leave one out estimate of  $\xi$  for the original data set, when one observation  $(x_i, Y_i)$  is removed, and  $\hat{\xi}_{(\cdot)} = \sum_{i=1}^N \hat{\xi}_{(-i)} / N$ . Notice that  $\pi_1 = \alpha/2$  and  $\pi_2 = 1 - \alpha/2$  when  $\hat{z} = \hat{a} = 0$ .

## Appendix E. Risks and odds ratios when interaction is removed

Suppose a specific model of no interaction  $M$  is chosen, and that we want to assess the amount of interaction among the co-factors in  $J$ . For a prospective study we use either of the three quantities  $AP(x; J, M)$ ,  $AP(J, M)$  or  $PAF(J, M)$ , and it follows from Table 5 that this requires expressions for  $\theta_{\text{rem}}(x, \psi; J, M)$ , the disease risk when interaction among the variables in  $J$  is absent. For a retrospective study we use either of the two quantities  $AP_{\text{OR}}(x; J, M)$  and  $AP_{\text{OR}}(J, M)$ , which requires analogous formulas for the odds ratio  $OR_{\text{rem}}(x, \psi; J, M)$  when there is no interaction among the variables in  $J$ . In this appendix we provide the required expressions for  $\theta_{\text{rem}}(x, \psi; J, M)$  and  $OR_{\text{rem}}(x, \psi; J, M)$ .



## E.1 Prospective studies

Table 6 provides expressions for  $\theta_{\text{rem}}(x, \psi; J, M)$  for some of the generalized linear models  $M$  of no interaction in eq. (17). Formulas are given for the linear and the two logistic parametrizations of  $\psi$  in eqs (22)–(24). Due to presence of the baseline risk parameter  $\psi_0$ , several expressions are quite complicated. For this reason, it is important to adapt parametrization to the assumed model of no interaction, using the

**Table 6:** Disease risk  $\theta_{\text{rem}}(x, \psi; J, M)$  for exposure profile  $x$  when interaction among the factors in  $J$  has been removed.

$p$	$M$	$x$	$J$	$\psi$	$\theta_{\text{rem}}(x, \psi; J, M)$
2	Mult	(1,1)	{1, 2}	lin	$\frac{\psi_{01}}{1-\psi_{01}} \cdot \frac{\psi_{10}}{1-\psi_{10}} \cdot \frac{1-\psi_{00}}{\psi_{00}} \left/ \left( 1 + \frac{\psi_{01}}{1-\psi_{01}} \cdot \frac{\psi_{10}}{1-\psi_{10}} \cdot \frac{1-\psi_{00}}{\psi_{00}} \right) \right.$
				logit	$\frac{\exp(\psi_{00} + \psi_{01} + \psi_{10})}{1 + \exp(\psi_{00} + \psi_{01} + \psi_{10})}$
	Add	(1,1)		lin	$\psi_{01} + \psi_{10} - \psi_{00}$
				logit	$\frac{\exp(\psi_{00} + \psi_{01})}{1 + \exp(\psi_{00} + \psi_{01})} + \frac{\exp(\psi_{00} + \psi_{10})}{1 + \exp(\psi_{00} + \psi_{10})} - \frac{\exp(\psi_{00})}{1 + \exp(\psi_{00})}$
	Disj	(1,1)		lin	$1 - (1 - \psi_{01})(1 - \psi_{10}) / (1 - \psi_{00})$
				logit	$1 - \exp[-\log(1 + e^{\psi_{00} + \psi_{01}}) - \log(1 + e^{\psi_{00} + \psi_{10}}) + \log(1 + e^{\psi_{00}})]$
3	Mult	(1,1,0)	{1, 2, 3}	lin	$\frac{\psi_{010}}{1-\psi_{010}} \cdot \frac{\psi_{100}}{1-\psi_{100}} \cdot \frac{1-\psi_{000}}{\psi_{000}} \left/ \left( 1 + \frac{\psi_{010}}{1-\psi_{010}} \cdot \frac{\psi_{100}}{1-\psi_{100}} \cdot \frac{1-\psi_{000}}{\psi_{000}} \right) \right.$
				logit	$\frac{\exp(\psi_{000} + \psi_{010} + \psi_{100})}{1 + \exp(\psi_{000} + \psi_{010} + \psi_{100})}$
		(1,0,1)		logit	$\frac{\exp(\psi_{000} + \psi_{001} + \psi_{100})}{1 + \exp(\psi_{000} + \psi_{001} + \psi_{100})}$
		(0,1,1)		logit	$\frac{\exp(\psi_{000} + \psi_{001} + \psi_{010})}{1 + \exp(\psi_{000} + \psi_{001} + \psi_{010})}$
		(1,1,1)		logit	$\frac{\exp(\psi_{000} + \psi_{001} + \psi_{010} + \psi_{100})}{1 + \exp(\psi_{000} + \psi_{001} + \psi_{010} + \psi_{100})}$
		(1,1,0)	{1, 2}	logit	$\frac{\exp(\psi_{000} + \psi_{010} + \psi_{100})}{1 + \exp(\psi_{000} + \psi_{010} + \psi_{100})}$
		(1,1,1)		[28]	$\frac{\exp(\psi_{000} + \psi_{001} + \psi_{010} + \psi_{100} + \psi_{011} + \psi_{101})}{1 + \exp(\psi_{000} + \psi_{001} + \psi_{010} + \psi_{100} + \psi_{011} + \psi_{101})}$
				[29]	$\frac{\exp(\psi_{000} + \psi_{011} + \psi_{101} - \psi_{001})}{1 + \exp(\psi_{000} + \psi_{011} + \psi_{101} - \psi_{001})}$
	Add	(1,1,0)	{1, 2, 3}	lin	$\psi_{010} + \psi_{100} - \psi_{000}$
				logit	$\frac{\exp(\psi_{000} + \psi_{010})}{1 + \exp(\psi_{000} + \psi_{010})} + \frac{\exp(\psi_{000} + \psi_{100})}{1 + \exp(\psi_{000} + \psi_{100})} - \frac{\exp(\psi_{000})}{1 + \exp(\psi_{000})}$
		(1,0,1)		lin	$\psi_{001} + \psi_{100} - \psi_{000}$
				logit	$\frac{\exp(\psi_{000} + \psi_{001})}{1 + \exp(\psi_{000} + \psi_{001})} + \frac{\exp(\psi_{000} + \psi_{100})}{1 + \exp(\psi_{000} + \psi_{100})} - \frac{\exp(\psi_{000})}{1 + \exp(\psi_{000})}$
		(0,1,1)		lin	$\psi_{001} + \psi_{010} - \psi_{000}$
				logit	$\frac{\exp(\psi_{000} + \psi_{001})}{1 + \exp(\psi_{000} + \psi_{001})} + \frac{\exp(\psi_{000} + \psi_{010})}{1 + \exp(\psi_{000} + \psi_{010})} - \frac{\exp(\psi_{000})}{1 + \exp(\psi_{000})}$
		(1,1,1)		lin	$\psi_{001} + \psi_{010} + \psi_{100} - 2\psi_{000}$
				logit	$\frac{\exp(\psi_{000} + \psi_{001})}{1 + \exp(\psi_{000} + \psi_{001})} + \frac{\exp(\psi_{000} + \psi_{010})}{1 + \exp(\psi_{000} + \psi_{010})} + \frac{\exp(\psi_{000} + \psi_{100})}{1 + \exp(\psi_{000} + \psi_{100})} - \frac{2\exp(\psi_{000})}{1 + \exp(\psi_{000})}$
		(1,1,0)	{1, 2}	lin	$\psi_{010} + \psi_{100} - \psi_{000}$
				logit	$\frac{\exp(\psi_{000} + \psi_{010})}{1 + \exp(\psi_{000} + \psi_{010})} + \frac{\exp(\psi_{000} + \psi_{100})}{1 + \exp(\psi_{000} + \psi_{100})} - \frac{\exp(\psi_{000})}{1 + \exp(\psi_{000})}$
		(1,1,1)		lin	$\psi_{011} + \psi_{101} - \psi_{001}$
				[28]	$\frac{\exp(\psi_{000} + \psi_{001} + \psi_{010} + \psi_{011})}{1 + \exp(\psi_{000} + \psi_{001} + \psi_{010} + \psi_{011})} + \frac{\exp(\psi_{000} + \psi_{001} + \psi_{100} + \psi_{101})}{1 + \exp(\psi_{000} + \psi_{001} + \psi_{100} + \psi_{101})} - \frac{\exp(\psi_{000} + \psi_{001})}{1 + \exp(\psi_{000} + \psi_{001})}$
				[29]	$\frac{\exp(\psi_{000} + \psi_{011})}{1 + \exp(\psi_{000} + \psi_{011})} + \frac{\exp(\psi_{000} + \psi_{101})}{1 + \exp(\psi_{000} + \psi_{101})} - \frac{\exp(\psi_{000} + \psi_{001})}{1 + \exp(\psi_{000} + \psi_{001})}$

Note: Each disease risk  $\theta_{\text{rem}}(x, \psi; J, M)$  is based on an assumed model of no interaction  $M$  among a subset  $J \subset \{1, \dots, p\}$  of the  $p$  co-factors, for a given exposure profile  $x = (x_1, \dots, x_p)$ . The vector  $\psi = (\psi_x)$  contains the  $2^p$  parameters, either on the linear [22] or logit risk scale [23]–[24]. For each model, the table includes those  $x$  for which at least two factors in  $J$  are turned on. An entry 'logit' in the parameter column means that both eqs (23) and (24) apply. All models  $M$  are from Table 1, with Mrisk = Multiplicative risk, Disj = Disjunctive, Add = Additive and Mult = Multiplicative.

linear parametrization for an additive model of no interaction, and any of the two logit parametrizations for a multiplicative model of no interaction. More generally, it is appropriate to parametrize  $\psi$  using the same link function transformation  $g$  of risk, as in eq. (17).

But our framework includes other ways of choosing the model of no interaction. When no variables are confounders ( $J = \{1, \dots, p\}$ ), two possible models are

$$\theta_{\text{rem}}(x, \psi; \{1, \dots, n\}, M) = \begin{cases} \psi_0 + \psi_1 1_{\{x \neq 0\}}, & M \text{ dominant,} \\ \psi_0 + \psi_1 1_{\{x=1\}}, & M \text{ recessive.} \end{cases} \quad (49)$$

They both involve two parameters  $\psi_0$  and  $\psi_1$  of the linear saturated model [22]; the disease risk for an individual exposed to no ( $x = 0 = (0, \dots, 0)$ ) or all ( $x = 1 = (1, \dots, 1)$ ) factors. It is also possible to divide the population into strata with different mechanisms of interaction (Westerlind et al. 2014), viewed as a Bayesian network (Koski and Noble 2009). Other models of no interaction impose monotonicity constraints on  $\theta_{\text{rem}}(x)$  (Traskin et al. 2013).

## E.2 Retrospective studies

Table 7 provides  $\text{OR}_{\text{rem}}(x, \psi; J, M)$  expressions for two models  $M$  of no interaction and either of the two logit parametrizations (25)–(26) of  $\psi$ . Since the baseline parameter is absent, many of the formulas in Table 7 are simpler than those of Table 6.

**Table 7:** Odds ratios  $\text{OR}_{\text{rem}}(x, \psi; J, M)$  for exposure profile  $x$  when interaction among the factors in  $J$  is removed.

$p$	$M$	$x$	$J$	$\psi$	$\text{OR}_{\text{rem}}(x, \psi; J, M)$
2	Multiplicative	(1,1)	{1, 2}	logit	$\exp(\psi_{01} + \psi_{10})$
	Additive odds	(1,1)		logit	$\exp(\psi_{01}) + \exp(\psi_{10}) - 1$
3	Multiplicative	(1,1,0)	{1, 2, 3}	logit	$\exp(\psi_{010} + \psi_{100})$
		(1,0,1)		logit	$\exp(\psi_{001} + \psi_{100})$
		(0,1,1)		logit	$\exp(\psi_{001} + \psi_{010})$
		(1,1,1)		logit	$\exp(\psi_{001} + \psi_{010} + \psi_{100})$
		(1,1,0)	{1, 2}	logit	$\exp(\psi_{010} + \psi_{100})$
		(1,1,1)		[30]	$\exp(\psi_{001} + \psi_{010} + \psi_{100} + \psi_{011} + \psi_{101})$
		(1,1,1)		[31]	$\exp(\psi_{011} + \psi_{101} - \psi_{001})$
	Additive odds	(1,1,0)	{1, 2, 3}	logit	$\exp(\psi_{010}) + \exp(\psi_{100}) - 1$
		(1,0,1)		logit	$\exp(\psi_{001}) + \exp(\psi_{100}) - 1$
		(0,1,1)		logit	$\exp(\psi_{001}) + \exp(\psi_{010}) - 1$
		(1,1,1)		logit	$\exp(\psi_{001}) + \exp(\psi_{010}) + \exp(\psi_{100}) - 2$
		(1,1,0)	{1, 2}	logit	$\exp(\psi_{010}) + \exp(\psi_{100}) - 1$
		(1,1,1)		[30]	$\exp(\psi_{010} + \psi_{001} + \psi_{011}) + \exp(\psi_{100} + \psi_{001} + \psi_{101}) - \exp(\psi_{001})$
		(1,1,1)		[31]	$\exp(\psi_{011}) + \exp(\psi_{101}) - \exp(\psi_{001})$

Note: The odds ratios  $\text{OR}_{\text{rem}}(x, \psi; J, M)$  are based on an assumed model  $M$  of no interaction among the factors in  $J \subset \{1, \dots, p\}$ , for a given exposure profile  $x$ , where  $\psi = (\psi_x; x \neq 0)$  contains the  $2^p - 1$  non-baseline parameters of the saturated logistic regression model [25] or [26]. An entry 'logit' for  $\psi$  means that both parametrizations (25) or (26) apply. For each model, the table includes those  $x$  for which at least two factors in  $J$  are turned on.

We can use Table 7 (or more generally formula (17)) to deduce that  $\text{AP}_{\text{OR}}(x; J)$  of joint effects is independent of the levels of the confounding variables when there is multiplicative non-interaction among the variables in  $J$ . Indeed, assuming an odds ratio parametrization (25), we notice that

$$\text{OR}(x, \psi) = \text{OR}_{\text{rem}}(x, \psi; J, \text{Multiplicative}) = \exp\left(\psi_{x_J} + \sum_{i \in J} x_i \psi_{x_J + e_i}\right). \quad (50)$$

Inserting eq. (50) into eq. (40), we find that

$$\text{AP}_{\text{OR}}(x; J) = \frac{O(x, \psi)/O(x_J, \psi) - 1}{\max[O(x, \psi)/O(x_J, \psi), 1]} = \frac{\exp(\sum_{i \in J} x_i \psi_{x_j + e_i}) - 1}{\max\left[\exp(\sum_{i \in J} x_i \psi_{x_j + e_i}), 1\right]}, \quad (51)$$

independently of the levels  $x_i$  of the confounding variables  $i \neq J$ .

## Appendix F. Details from simulation study

The retrospective simulation study of Section 6 (and Table 3) contains a number of data sets with  $N/2$  cases and  $N/2$  controls, for three models I, II and III of no interaction. For each simulated data set, exposure profiles were generated independently from

$$P(x|Y=y) = \begin{cases} q(x)[1 - \theta(x)]/[1 - E(\theta(X))], & y = 0, \\ q(x)\theta(x)/E(\theta(X)), & y = 1, \end{cases} \quad (52)$$

and some conveniently chosen prior distribution  $q$ . In most runs of the present simulation study, the prior eq. (52) was uniform  $q(x) \equiv 2^{-p}$ .

If  $N$  is too small, estimates of  $\psi$  may be ill-conditioned for data sets that lack some exposure profile  $x$  among cases ( $N_{x1} = 0$ ) or among controls ( $N_{x0} = 0$ ). This may happen if either  $\theta(x)$  or  $1 - \theta(x)$  is small. Since the saturated model is used for inference,  $N \gg 2^p$  is required. For simulated data sets, the prior distribution  $q$  in eq. (52) can be adjusted in order to minimize the required  $N$ . This is particularly important for the bootstrap confidence interval method of Appendix D, since each simulated data set must be resampled  $B$  times, and some of the resampled estimates  $\hat{\psi}^*$  may be even more ill-conditioned than  $\hat{\psi}$ .

In order to avoid ill-conditioned estimates, we employed non-uniform priors in Table 3 for the following runs: For Model I we used  $q(0, 0) = 4/7$ ,  $q(1, 0) = q(0, 1) = q(1, 1) = 1/7$  for all three methods when  $N = 500$ , and for the bootstrap also when  $N = 1000$ . For Model II we chose  $q(0, 0) = 2/7$ ,  $q(1, 0) = 3/7$ ,  $q(0, 1) = q(1, 1) = 1/7$  for all three methods when  $N = 500$ , and for the bootstrap also when  $N = 1000$ . For Model III we defined  $q(x)$  to be inversely proportional to  $\min[\theta(x), 1 - \theta(x)]$  for the bootstrap when  $N = 1000$ .

**Acknowledgements:** Support of Ola Hössjer from the Swedish Research Council, contract 621-2013-4633 is acknowledged, as well as valuable comments from an associate editor and a reviewer.

## References

- Agresti, A. (2013). *Categorical Data Analysis*. 3rd edition. Hoboken, NJ: Wiley.
- Andersson, T., Alfredsson, L., Källberg, H., Zdravkovic, S. and Ahlbom, A. (2005). Calculating measures of biological interaction. *European Journal of Epidemiology*, 20:575–579.
- Aschengrau, A. and Seage, G. R. (2003). *Essentials of Epidemiology in Public Health*. Epidemiology Series. Burlington, MA: Jones & Bartlett Publishers.
- Assman, S. F., Hosmer, D. W., Lemeshow, S. and Mundt, K. A. (1996). Confidence intervals for measures of interaction. *Epidemiology*, 7(3):286–290.
- Benichou, J. (2001). A review of adjusted estimators of attributable risk. *Statistical Methods in Medical Research*, 10:195–216.
- Bookman, E. B., McAllister, K., Gillanders, E., Wanke, K., Balshaw, D., Rutter, J., Reedy, J., Shaughnessy, D., Agurs-Collins, T., Paltoo, D., Atienza, A., Bierut, L., Kraft, P., Fallin, M. D., Perera, F., Turkheimer, E., Boardman, J., Marazita, M. L., Rappaport, S. M., Boerwinkle, E., Suomi, S. J., Caporaso, N. E., Hertz-Picciotto, I., Jacobson, K. C., Lowe, W. L., Goldman, L. R., Duggal, P., Gunnar, M. R., Manolio, T. A., Green, E. D., Olster, D. H., Birnbaum, L. S., for the NIH GxE Interplay Workshop Participants. (2011). Gene-environment interplay in common complex diseases: Forging an integrative model – recommendations from an NIH workshop. *Genetic Epidemiology*, 35:217–225.

- Cordell, H. J. (2002). Epistasis: What it means, what it doesn't mean, and statistical methods to detect it in humans. *Human Molecular Genetics*, 11:2463–2468.
- Cornfield, J. (1951). A method of estimating comparative rates from clinical data. Application to cancer of the lung, breast and cervix. *Journal of the National Cancer Institute*, 11:1269–1275.
- Cree, B. (2014). Multiple sclerosis genetics. *Handbook of Clinical Neurology*, 122:193–209.
- Deubner, D. C., Wilkinson, W. E., Helms, M. J., Tyroler, H. A. and Hames, C. G. (1980). Logistic model estimation of death attributable to risk factors for cardiovascular disease in Evans County, Georgia. *American Journal of Epidemiology*, 112:135–143.
- Efron, B. (1987). Better bootstrap confidence intervals (with discussion). *Journal of the American Statistical Association*, 82:171–200.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability 57. Boca Raton, Florida: Chapman and Hall/CRC.
- Eichler, E. E., Flint, J., Gibson, G., Kong, A., Leal, S. M., Moore, J. H. and Nadeau, J. H. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews Genetics*, 11(6):446–450.
- Eide, G. E. (2001). Attributable fractions: Fundamental concepts and their visualization. *Statistical Methods in Medical Research*, 10:159–193.
- Greenland, S. (1983). Tests for interaction in epidemiological studies: A review and a study of power. *Statistics in Medicine*, 2(2):243–251.
- Greenland, S. (1993). Additive risk versus additive relative risk models. *Epidemiology*, 4(1):32–36.
- Greenland, S. and Robins, J. (1988). Conceptual problems in the definition and interpretation of attributable fractions. *American Journal of Epidemiology*, 128:1185–1197.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. New York: Wiley.
- Hedström, A. K., Sundqvist, E., Bäärnhielm, M., Nordin, N., Hillert, J., Kockum, I., Olsson, T. and Alfredsson, L. (2011). Smoking and two human leukocyte antigen genes interact to increase the risk for multiple sclerosis. *Brain*, 134:653–664.
- Hosmer, D. W. and Lemeshow, S. (1992). Confidence interval estimation of interaction. *Epidemiology*, 3(5):452–456.
- Kalilani, L. and Atashili, J. (2006). Measuring additive interaction using odds ratios. *Epidemiologic Perspectives & Innovations*, 3:5.
- Knol, M. J., VanderWeele, T. J., Groenwold, R. H. H., Klungel, O. H., Rovers, M. M. and Grobbee, D. E. (2011). Estimating measures of interaction on an additive scale for preventive exposures. *European Journal of Epidemiology*, 26:433–438.
- Koski, T. and Noble, J. (2009). *Bayesian Networks: An Introduction*. Chichester, United Kingdom: John Wiley and Sons.
- Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*. New York: Springer.
- Lekman, M., Hössjer O., Andrews, P., Källberg, H., Uvehag, D., Charney, D., Manji, H., Rush, J. A., McMahon, F. J., Moore, J. H. and Kockum, I. (2014). The genetic interacting landscape of 63 candidate genes in major depressive disorder: An explorative study. *Biodata Mining*, 7:19.
- Levin, M. L. (1953). The occurrence of lung cancer in man. *Acta – Unio Internationalis Contra Cancrum*, 9(3):531–541.
- Maher, B. (2008). The case of the missing heritability. *Nature*, 456(6):18–21.
- McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*. 2nd edition. London: Chapman & Hall.
- Miettinen, O. S. (1976). Estimability and estimation in case-referent studies. *American Journal of Epidemiology*, 103:226–235.
- Moore, J. H. (2004). Computational analysis of gene-gene interactions using multifactor dimensionality reduction. *Expert Review of Molecular Diagnostics*, 4:795–803.
- Moore, J. H., Gilbert, J. C., Tsai, C. -T., Chiang, F. -T., Holden, T., Barney, N. and White, B. C. (2006). A flexible computational framework for detecting, characterizing and detecting statistical patterns of epistasis in genetic studies of human disease susceptibility. *Journal of Theoretical Biology*, 241:252–261.
- Northridge, M. E. (1995). Public health methods- attributable risk as a link between causality and public health action. *American Journal of Public Health*, 85(9):1202–1203.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.
- Phillips, P. C. (2008). Epistasis – the essential role of genetic interactions in the structure of evolution of genetic systems. *Nature Reviews Genetics*, 9:855–867.
- Prentice, R. L. and Pyke, R. (1979). Logistic disease incidence models and case-control studies. *Biometrika*, 66(3):403–411.
- Richardson, D. B. and Kaufman, J. S. (2009). Estimation of the relative excess risk due to interaction and associated confidence intervals. *American Journal of Epidemiology*, 169:756–760.
- Rothman, K. J. (1976a). Causes. *American Journal of Epidemiology*, 104:587–592.
- Rothman, K. J. (1976b). The estimation of synergy or antagonism. *American Journal of Epidemiology*, 103:506–511.
- Rothman, K. J. (2012). *Epidemiology. An introduction*. 2nd edition. NY: Oxford Univ. Press.
- Sjölander, A., Lee, W., Källberg, H. and Pawitan, Y. (2014). Bounds on causal interaction for binary outcomes. *Biometrics*, 70:500–505.

- Sjölander, A. and Vansteelandt, S. (2011). Doubly robust estimation of attributable fractions. *Biostatistics*, 12:112–121.
- Skrondal, A. (2003). Interaction as departure from additivity in case-control studies: A cautionary note. *American Journal of Epidemiology*, 158:251–258.
- Traskin, M., Wang, W., Ten Have, T. R. and Small, D. S. (2013). Efficient estimation of the attributable fraction when there are monotonicity constraints and interactions. *Biostatistics*, 14(1):173–188.
- VanderWeele, T. J. (2009). Sufficient cause interactions and statistical interactions. *Epidemiology*, 20(1):6–13.
- VanderWeele, T. J. (2010). Epistatic interactions. *Statistical Applications in Genetics and Molecular Biology*, 9(1):Article 1.
- VanderWeele, T. J. (2013). Reconsidering the denominator of the attributable proportion for interaction. *European Journal of Epidemiology*, 28:779–784.
- VanderWeele, T. J. and Knol, M. (2014). A tutorial on interaction. *Epidemiologic Methods*, 3(1):33–72.
- VanderWeele, T. J. and Robins, J. M. (2008). Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika*, 95:49–61.
- Westerlund, H., Jääskinen, V., Corander, J., Hillert, J. and Koski, T. (2014). The learning for mixtures of multicausal interaction networks. In *Westerlund: Modeling genetic susceptibility to Multiple Sclerosis*, PhD thesis, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm.
- Zou, G. Y. (2008). On the estimation of additive interaction by use of the four-by-two table and beyond. *American Journal of Epidemiology*, 168(2):212–224.