

Quantifying and estimating additive measures of interaction from case-control data

Ola Hössjer^{a,*}, Lars Alfredsson^b, Anna Karin Hedström^b, Magnus Lekman^c,
Ingrid Kockum^c, Tomas Olsson^c

^a*Department of Mathematics, Stockholm University, Stockholm, Sweden*

^b*Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden*

^c*Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden*

ola@math.su.se (O. Hössjer), Lars.Alfredsson@ki.se (L. Alfredsson),
anna.hedstrom@ki.se (A. Hedström), magnus.lekman@gmail.com (M. Lekman),
Ingrid.Kockum@ki.se (I. Kockum), Tomas.Olsson@ki.se (T. Olsson)

Received: 22 March 2017, Revised: 12 April 2017, Accepted: 12 April 2017,
Published online: 26 April 2017

Abstract In this paper we develop a general framework for quantifying how binary risk factors jointly influence a binary outcome. Our key result is an additive expansion of odds ratios as a sum of marginal effects and interaction terms of varying order. These odds ratio expansions are used for estimating the excess odds ratio, attributable proportion and synergy index for a case-control dataset by means of maximum likelihood from a logistic regression model. The confidence intervals associated with these estimates of joint effects and interaction of risk factors rely on the delta method. Our methodology is illustrated with a large Nordic meta dataset for multiple sclerosis. It combines four studies, with a total of 6265 cases and 8401 controls. It has three risk factors (smoking and two genetic factors) and a number of other confounding variables.

Keywords Additive odds model, attributable proportion, case-control data, expansion of odds ratios, interaction of risk factors, logistic regression

2010 MSC [62F10](#), [62F12](#), [62F25](#), [62J12](#), [62P10](#)

*Corresponding author.

1 Introduction

Many complex diseases are influenced by a number of risk factors that interact in a complicated way. This is often quantified by means of a regression model with affection status of a given disease as binary response, whereas the risk factors and possibly some other variables are chosen as covariates. Logistic regression models have often been used to quantify main effects and strength of interaction among the risk factors with regards to disease. There are several reasons for this. The logistic transformation is first of all the canonical link of a generalized linear model with a binomially distributed response, and the parameters of this model have a straightforward multiplicative odds interpretation [20]. A second reason is that many epidemiological datasets are collected retrospectively based on outcomes rather than on covariates. In particular, it is well known that many parameters of the logistic regression model can be estimated consistently for case-control studies under suitable sampling assumptions on the cases and controls [24]. There are also additive models of joint effects and interaction. They have recently gained in popularity, since they are believed to approximate a biological system with causal mechanisms more accurately than multiplicative odds [30, 26]. The additive measures of interaction are functions of relative risks between individuals with or without exposure to the risk factors, such as the relative excess of risk due to interaction (RERI), the attributable proportion (AP) due to interaction or the synergy index (SI) [12, 28]. Although these relative risks cannot be estimated consistently for case-control data, they are well approximated by estimable odds ratios when the disease risk is small [15, 4, 3]. When relative risks are replaced by odds ratios in the expressions for RERI, AP and SI, they correspond to additive odds models of interaction.

In previous work [16] we developed a unified theory for quantifying and estimating different measures of marginal effects, joint effects and interaction among risk factors on a multiplicative, additive, additive odds or some other scale. We also described how to estimate and produce confidence intervals for these quantities from a prospective study, with data sampled based on their covariates, or from a case-control study. Traditional definitions of attributable proportion have an unbounded negative range [18, 29, 19]. In [16], we introduced a novel normalization of AP that guarantees a range between -1 and 1 , with negative or positive values depending on whether there is synergism or antagonism between the risk factors.

In this paper we concentrate on case-control data and additive odds measures of main effects, joint effects and interaction. Our main result is to express the odds ratio of the risk factors as a sum of terms, which include their main effects and different orders of interaction, when the effect of other confounding covariates is controlled for. In this way we extend and unify some previously used measures of interaction [23, 17, 16]. In order to find confidence intervals (CIs) for the attributable proportion, synergy index and excess odds ratio due to joint effects and interaction from case-control data, we first estimate the parameters of a logistic regression model by maximum likelihood. Then we use the delta method [27, 6] for an appropriate transformation of these measures of joint effects and interaction in order to find their standard errors and CIs on the new scale, before transforming back to the original odds scale.

The paper is organized as follows: In Section 2 we define a logistic regression model that includes marginal and interaction effects for the risk factors of interest,

as well as marginal effects of other confounding covariates. The additive odds measures of joint effects and interaction are defined in Section 3, and the procedures for estimating them are given in Section 4. Our methodology is illustrated in Section 5 for a multiple sclerosis (MS) dataset from Hedström et al. [14], and a concluding discussion appears in Section 6.

2 A logistic regression model

Let $Y \in \{0, 1\}$ be a binary outcome variable, with $Y = 1$ for individuals that carry a certain disease, and $Y = 0$ for those that do not. Consider a large population, and let

$$\theta_{\mathbf{x}} = P(Y = 1 | \mathbf{x}) \quad (1)$$

be the disease probability of a randomly chosen individual. It is assumed to be a function of $p + q$ covariates $\mathbf{x} = (\mathbf{v}, \mathbf{z}) = (v_1, \dots, v_p, z_1, \dots, z_q)$, of which the first p are binary risk factors v_1, \dots, v_p of main interest, with $v_j = 1$ indicating presence and $v_j = 0$ absence of each such factor j . The other q covariates z_1, \dots, z_q are not necessarily binary, and they are included in the model as possible confounders. We will parametrize the disease probability on a logit scale

$$\text{logit } \theta_{\mathbf{x}} = \log \frac{\theta_{\mathbf{x}}}{1 - \theta_{\mathbf{x}}} = \kappa_0 + \sum_{\mathbf{0} < \mathbf{w} \leq \mathbf{v}} \psi_{\mathbf{w}} + \sum_{j=1}^q \kappa_j z_j, \quad (2)$$

with $\mathbf{w} = (w_1, \dots, w_p)$ a binary vector and $\mathbf{0} = (0, \dots, 0)$ a vector with zero components. Inequalities between vectors are interpreted componentwise, so that the first sum in (2) is taken over all nonzero vectors \mathbf{w} such that $w_j \leq v_j$, with $w_j = 1$ for at least one factor j . This sum equals 0 when $\mathbf{v} = \mathbf{0}$.

The logistic model in (2) is saturated for the binary risk factors, including all orders of their interaction, whereas it is linear in the other covariates [1, 2]. In particular, if $v_j = 1$ and all other components of \mathbf{v} are zero, $\psi_{\mathbf{v}}$ quantifies the marginal effect of factor j in absence of the others, whereas $\psi_{\mathbf{v}}$ for \mathbf{v} such that $|\mathbf{v}| = \sum_j v_j \geq 2$ quantifies interaction of order $|\mathbf{v}|$ among those factors j for which $v_j = 1$.

It is assumed that $\boldsymbol{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_q)$ are nuisance parameters, whereas

$$\boldsymbol{\psi} = (\psi_{\mathbf{v}}; \mathbf{0} < \mathbf{v} \leq \mathbf{1}) \quad (3)$$

are the $2^p - 1$ structural parameters of main interest that quantify marginal effects and interaction among the p risk factors, with $\mathbf{1} = (1, \dots, 1)$. Since they are defined on a logit scale, it is possible to estimate them from suitable case-control datasets. In the next section we will develop alternative measures of marginal effects, joint effects and interaction that are functions of $\boldsymbol{\psi}$, but expressed in terms of odds ratios.

3 Measures of joint effects and interaction

Our purpose is to estimate and produce confidence intervals for parameters

$$\xi = \xi(\boldsymbol{\psi}) \quad (4)$$

that quantify marginal effects, joint effects or interaction between a subset $J \subset \{1, \dots, p\}$ of factors, when the levels of the remaining factors in $K = \{1, \dots, p\} \setminus J$,

and the q other confounder variables are controlled for. It is assumed that (4) only involves the structural parameters ψ , not the nuisance parameters κ . The reason is that we consider parameters (4) that are functions of odds ratios

$$\text{OR}_{\mathbf{v}} = \frac{\theta(\mathbf{v}, \mathbf{z}) / (1 - \theta(\mathbf{v}, \mathbf{z}))}{\theta(\mathbf{0}, \mathbf{z}) / (1 - \theta(\mathbf{0}, \mathbf{z}))} = \exp\left(\sum_{\mathbf{0} < \mathbf{w} \leq \mathbf{v}} \psi_{\mathbf{w}}\right), \tag{5}$$

i.e. ratios between the disease odds of two subjects with identical confounding variables \mathbf{z} but different risk exposures \mathbf{v} and $\mathbf{0}$ respectively, for some or all of the $2^p - 1$ vectors $\mathbf{0} < \mathbf{v} \leq \mathbf{1}$. In particular, when the disease risks in (1) are small, the odds ratio in (5) approximates the relative risk

$$\text{RR}_{(\mathbf{v}, \mathbf{z})} = \frac{\theta(\mathbf{v}, \mathbf{z})}{\theta(\mathbf{0}, \mathbf{z})} \tag{6}$$

well.

After possible reordering of factors we may assume without loss of generality that those in J come first, so that $\mathbf{v} = (\mathbf{v}_J, \mathbf{v}_K)$, with $\mathbf{v}_J = (v_j; j \in J)$ and $\mathbf{v}_K = (v_j; j \in K)$ being the exposure levels of the factors in J and K respectively. We will consider linear combinations of the odds ratios (5), and the following concept will be central for the rest of the paper:

Definition 1. Suppose the exposure levels \mathbf{v}_K of the confounding risk factors are fixed. For any \mathbf{v}_J we introduce the additive increment

$$\Delta^{(\mathbf{v}_J, \mathbf{v}_K)} \text{OR} = \sum_{\mathbf{0} \leq \mathbf{w} \leq \mathbf{v}_J} (-1)^{|\mathbf{v}_J - \mathbf{w}|} \text{OR}_{(\mathbf{w}, \mathbf{v}_K)} \tag{7}$$

of the odds ratio of order \mathbf{v}_J , where summation on the right-hand side of (7) is over all binary vectors $\mathbf{w} = (w_j; j \in J)$ of length $|J|$ whose coordinates do not exceed those of \mathbf{v}_J .

When $|\mathbf{v}_J| = 0$, (7) is the odds ratio when none of the risk variables in J are turned on, and when $|\mathbf{v}_J| = 1$ it is the marginal odds ratio increment of one single factor. For $|\mathbf{v}_J| \geq 2$ we interpret (7) as a measure of additive odds interaction among those factors $\{j \in J; v_j = 1\}$ in J that are at risk, when all interaction terms of lower order among these factors have been removed. More explicitly, the terms of order 0, 1, 2 and 3 in (7) have the form

$$\Delta^{(\mathbf{v}_J, \mathbf{v}_K)} \text{OR} = \begin{cases} \text{OR}_{(\mathbf{0}, \mathbf{v}_K), & \mathbf{v}_J = \mathbf{0}, \\ \text{OR}_{(\mathbf{e}_j, \mathbf{v}_K)} - \text{OR}_{(\mathbf{0}, \mathbf{v}_K)}, & \mathbf{v}_J = \mathbf{e}_j, \\ \text{OR}_{(\mathbf{e}_{jk}, \mathbf{v}_K)} - \text{OR}_{(\mathbf{e}_j, \mathbf{v}_K)} - \text{OR}_{(\mathbf{e}_k, \mathbf{v}_K)} + \text{OR}_{(\mathbf{0}, \mathbf{v}_K)}, & \mathbf{v}_J = \mathbf{e}_{jk}, \\ \text{OR}_{(\mathbf{e}_{jkl}, \mathbf{v}_K)} - \text{OR}_{(\mathbf{e}_{jk}, \mathbf{v}_K)} - \text{OR}_{(\mathbf{e}_{jl}, \mathbf{v}_K)} \\ \quad - \text{OR}_{(\mathbf{e}_{kl}, \mathbf{v}_K)} + \text{OR}_{(\mathbf{e}_j, \mathbf{v}_K)} + \text{OR}_{(\mathbf{e}_k, \mathbf{v}_K)} \\ \quad + \text{OR}_{(\mathbf{e}_l, \mathbf{v}_K)} - \text{OR}_{(\mathbf{0}, \mathbf{v}_K)}, & \mathbf{v}_J = \mathbf{e}_{jkl}, \end{cases} \tag{8}$$

where \mathbf{e}_j , \mathbf{e}_{jk} and \mathbf{e}_{jkl} are binary vectors of length $|J|$ with zero components, except in positions $\{j\}$, $\{j, k\}$ and $\{j, k, l\}$ respectively.

It is possible to give another interpretation of (7). Let $\mathbf{w} = (w_j; j \in J)$ be a vector with $0 \leq w_j \leq 1$, and extend the domain of the first $|J|$ components of the odds ratio $OR_{(\mathbf{w}, \mathbf{v}_K)}$ to $[0, 1]^{|J|}$. If \mathbf{v}_J is a binary vector of length $|\mathbf{v}_J| = l > 0$ with nonzero components $v_{j_1} = \dots = v_{j_l} = 1$, then $\Delta^{(\mathbf{v}_J, \mathbf{v}_K)}OR$ is a finite difference approximation of $\partial^l OR_{(\mathbf{w}, \mathbf{v}_K)} / (\partial w_{j_1} \dots \partial w_{j_l})$, where $\mathbf{w} = \mathbf{v}_J/2$.

Having defined the odds ratio increment in (7), we are ready to state the main result. It tells that the odds ratio for exposure \mathbf{v}_J is a sum of the odds ratio increments. This includes the baseline odds ratio when none of the factors in J are turned on, the additive marginal odds ratio increments for those $|\mathbf{v}_J|$ factors in J that are exposed, and the odds ratio interactions of order 2, 3, \dots , $|\mathbf{v}_J|$ among these factors (see Appendix A for a proof):

Theorem 1. *Suppose \mathbf{v}_K is fixed. Then the odds ratios $OR_{\mathbf{v}}$ admit an additive expansion*

$$OR_{\mathbf{v}} = OR_{(\mathbf{v}_J, \mathbf{v}_K)} = \sum_{\mathbf{0} \leq \mathbf{w} \leq \mathbf{v}_J} \Delta^{(\mathbf{w}, \mathbf{v}_K)}OR \quad (9)$$

for all \mathbf{v}_J , with $\Delta^{(\mathbf{w}, \mathbf{v}_K)}OR$ as defined in (7). Conversely, if (9) holds and $\Delta^{(\mathbf{w}, \mathbf{v}_K)}OR$ is a linear combination of $\{OR_{(\mathbf{u}, \mathbf{v}_K)}; \mathbf{0} \leq \mathbf{u} \leq \mathbf{w}\}$ for each $\mathbf{0} \leq \mathbf{w} \leq \mathbf{v}_J$, then necessarily $\Delta^{(\mathbf{w}, \mathbf{v}_K)}OR$ must satisfy (7).

The expansions in (7) and (9) define a combinatorial inclusion–exclusion principle. In order to see this, we associate to each $j \in J$ a set A_j and its complement A_j^c . Then $\Delta^{(\mathbf{v}_J, \mathbf{v}_K)}OR$ represents an intersection $(\bigcap_{j; v_j=1} A_j) \cap (\bigcap_{j; v_j=0} A_j^c)$, whereas $OR_{(\mathbf{v}_J, \mathbf{v}_K)}$ corresponds to a union $\bigcup_{j; v_j=1} A_j$, and in particular $OR_{(\mathbf{0}, \mathbf{v}_K)} = \emptyset$.

It is of interest to know how much of the odds ratio (5) can be explained by marginal effects and lower order interaction among the factors in J . For this reason we introduce the following concept:

Definition 2. A prediction

$$OR_{(\mathbf{v}_J, \mathbf{v}_K), i} = \sum_{\substack{\mathbf{0} \leq \mathbf{w} \leq \mathbf{v}_J \\ |\mathbf{w}| \leq i}} \Delta^{(\mathbf{w}, \mathbf{v}_K)}OR \quad (10)$$

of the odds ratio (5) is obtained by including only terms up to order i in (9), for some $0 \leq i \leq |\mathbf{v}_J|$.

It is possible to rewrite the predicted odds ratio (10) as a linear combination of lower order odds ratios $\{OR_{(\mathbf{w}, \mathbf{v}_K)}; \mathbf{0} \leq \mathbf{w} \leq \mathbf{v}_J, |\mathbf{w}| \leq i\}$ for exposure vectors \mathbf{w} that have at most i factors in J at risk (see Appendix B for a proof):

Proposition 1. *The prediction of the odds ratio $OR_{(\mathbf{v}_J, \mathbf{v}_K)}$ in (10), based on marginal effects among the factors in J at risk and their interaction terms up to order i , satisfies*

$$OR_{(\mathbf{v}_J, \mathbf{v}_K), i} = \sum_{\substack{\mathbf{0} \leq \mathbf{w} \leq \mathbf{v}_J \\ |\mathbf{w}| \leq i}} OR_{(\mathbf{w}, \mathbf{v}_K)} (-1)^{i-|\mathbf{w}|} \binom{|\mathbf{v}_J| - 1 - |\mathbf{w}|}{i - |\mathbf{w}|} \quad (11)$$

if $0 \leq i < |\mathbf{v}_J|$, and

$$OR_{(\mathbf{v}_J, \mathbf{v}_K), i} = OR_{(\mathbf{v}_J, \mathbf{v}_K)} \quad (12)$$

if $i = |\mathbf{v}_J|$.

Assume from now on that $v_J = \mathbf{1}$, so that all factors in J are at risk. We will quantify how much of the odds ratio $\text{OR}_{(\mathbf{1}, v_K)}$ is left unexplained by marginal effects and lower orders of interaction among the factors in J . To this end, we introduce the following concept:

Definition 3. The unadjusted excess odds ratio

$$\text{UnadjEOR}_{(\mathbf{1}, v_K), i} = \text{OR}_{(\mathbf{1}, v_K)} - \text{OR}_{(\mathbf{1}, v_K), i-1} \tag{13}$$

is the difference between the odds ratio (5) and a prediction (10) of it due to terms of order less than i , where $1 \leq i \leq |J|$.

The unadjusted excess odds ratio can be interpreted as an unstandardized residual of a regression model, where only marginal effects and interaction terms up to order $i - 1$ are included as independent variables in order to predict the odds ratio (5). Equivalently, it quantifies the contribution of the odds ratio expansion (9) from terms of order at least i . The special case $i = |J|$ is treated by Katsoulis and Bamia [17] when $K = \emptyset$. The definition of the unadjusted odds ratio then simplifies to $\Delta^{(\mathbf{1}, v_K)}\text{OR}$.

We will define three measures (4) of marginal, joint or interaction effects among the factors in J , and all of them are functions of UnadjEOR .

Definition 4. The excess odds ratio

$$\begin{aligned} \xi &= \text{EOR}_{(\mathbf{1}, v_K), i} \\ &= \frac{\text{OR}_{(\mathbf{1}, v_K)} - \text{OR}_{(\mathbf{1}, v_K), i-1}}{\text{OR}_{(\mathbf{0}, v_K)} - \text{OR}_{(\mathbf{0}, v_K)}} \\ &= \frac{\text{UnadjEOR}_{(\mathbf{1}, v_K), i}}{\text{OR}_{(\mathbf{0}, v_K)}} \end{aligned} \tag{14}$$

expresses (13) in units of the odds ratio when no factor in J is at risk, but those in K are kept fixed at level v_K . It has a range of $(-\infty, \infty)$, with $\text{EOR} = 0$ indicating absence of effect.

Definition 5. The quantity

$$\begin{aligned} \xi &= \text{AP}_{(\mathbf{1}, v_K), i} \\ &= \frac{\text{UnadjEOR}_{(\mathbf{1}, v_K), i}}{\max(\text{OR}_{(\mathbf{1}, v_K)}, \text{OR}_{(\mathbf{1}, v_K), i-1})} \\ &= \frac{\text{UnadjEOR}_{(\mathbf{1}, v_K), i}}{\max(\text{OR}_{(\mathbf{1}, v_K)}, \text{OR}_{(\mathbf{1}, v_K)} - \text{UnadjEOR}_{(\mathbf{1}, v_K), i})} \end{aligned} \tag{15}$$

is the attributable proportion of the odds ratio due to terms (7) of order at least i . It uses the same denominator as in Hössjer et al. [16] in order to assure $-1 \leq \text{AP} \leq 1$, with $\text{AP} = 0$ indicating absence of effect.

Definition 6. The synergy index

$$\begin{aligned} \xi &= \text{SI}_{(\mathbf{1}, v_I), i} \\ &= \frac{\text{OR}_{(\mathbf{1}, v_K)} - \text{OR}_{(\mathbf{0}, v_K)}}{\text{OR}_{(\mathbf{1}, v_K), i-1} - \text{OR}_{(\mathbf{0}, v_K)}} \end{aligned}$$

$$= \frac{\text{OR}_{(\mathbf{1}, \mathbf{v}_K)} - \text{OR}_{(\mathbf{0}, \mathbf{v}_K)}}{\text{OR}_{(\mathbf{1}, \mathbf{v}_K)} - \text{UnadjEOR}_{(\mathbf{1}, \mathbf{v}_K), i} - \text{OR}_{(\mathbf{0}, \mathbf{v}_K)}} \quad (16)$$

is only defined in order to quantify interaction ($2 \leq i \leq |J|$), since otherwise the denominator of (16) vanishes. It is also required that the joint and lower order effects of the factors in J are positive ($\text{OR}_{(\mathbf{1}, \mathbf{v}_K)} > \text{OR}_{(\mathbf{0}, \mathbf{v}_K)}$ and $\text{OR}_{(\mathbf{1}, \mathbf{v}_K), i-1} > \text{OR}_{(\mathbf{0}, \mathbf{v}_K)}$) in order for SI to be meaningful. Then its range is $(0, \infty)$, with $\text{SI} = 1$ indicating absence of effect.

All three quantities in equations (14)–(16) are stratified for the risk factors in K , like confounders that are controlled at their observed levels. In contrast, there is only partial (additive) control for the remaining q covariates of z . EOR and AP can be viewed as different types of standardized residuals of a regression model where marginal effects and lower order interactions among the factors in J are used to predict the odds ratio (5). The inverse synergy index SI^{-1} is the analogue of the coefficient of determination. It quantifies how large fraction of the odds ratio increment above the baseline level $\text{OR}_{(\mathbf{0}, \mathbf{v}_K)}$ that is explained by the regression model.

Some special cases of formulas (14)–(16) are of particular interest. When $i = |J| = 1$, there is one single risk factor ($J = \{1\}$). Equation (14) then quantifies the excess odds ratio due the marginal effect (EORM) of factor 1, when those in K are controlled for at level \mathbf{v}_K . Similarly, equation (15) defines the attributable proportion of risk due to the marginal effect of factor 1, whereas the synergy index is not well defined.

When $i = 1$ and $|J| \geq 2$, we refer to the quantity in (14) as EORJ, the excess odds ratio due to joint (marginal and interaction) effects of all factors in J when those in K are controlled at level \mathbf{v}_K . In the same way, AP is the attributable proportion due to joint effects among the factors in J , whereas the synergy index is undefined.

When $i = 2$ we refer to the quantity in (14) as the total excess odds ratio due to all levels of interaction (TotEORI) among the factors in J , when those in K are controlled at level \mathbf{v}_K . AP and SI are similarly referred to as the attributable proportion and synergy index due to all levels of interaction among the factors in J .

Finally, when $i = |J|$ we refer to the quantity in (14) as the excess odds ratio due to the highest order $|J|$ of interaction (EORI) among the factors in J , when those in K are controlled at level \mathbf{v}_K . Analogously, equations (15)–(16) quantify the attributable proportion and the synergy index due to the highest order $|J|$ of interaction among the factors in J .

Since EOR, AP and SI are all functions of UnadjEOR, it suffices to specify the latter. In the subsections to follow we will do so for models with 1, 2, or 3 risk factors in J .

3.1 One risk factor

When $|J| = 1$ there is only one possible unadjusted excess odds ratio

$$\text{UnadjEOR}_{(\mathbf{1}, \mathbf{v}_K), 1} = \text{OR}_{(\mathbf{1}, \mathbf{v}_K)} - \text{OR}_{(\mathbf{0}, \mathbf{v}_K)},$$

caused by the marginal effect of factor 1.

3.2 Two risk factors

When $|J| = 2$ there are only two possible unadjusted excess odds ratios

$$\text{UnadjEOR}_{(1,1,v_K),i} = \begin{cases} \text{OR}_{(1,1,v_K)} - \text{OR}_{(0,0,v_K)}, & i = 1, \\ \text{OR}_{(1,1,v_K)} - \text{OR}_{(1,0,v_K)} - \text{OR}_{(0,1,v_K)} + \text{OR}_{(0,0,v_K)}, & i = 2, \end{cases}$$

obtained by inserting (10) into (13). The first unadjusted excess odds ratio is due to the joint (marginal and interaction) effect of factors 1 and 2, whereas the second only includes interaction between these two factors.

3.3 Three risk factors

When $|J| = 3$ there are three possible unadjusted excess odds ratios

$$\text{UnadjEOR}_{(1,1,1,v_K),i} = \begin{cases} \text{OR}_{(1,1,1,v_K)} - \text{OR}_{(0,0,0,v_K)}, & i = 1, \\ \text{OR}_{(1,1,1,v_K)} - \text{OR}_{(1,0,0,v_K)} \\ \quad - \text{OR}_{(0,1,0,v_K)} - \text{OR}_{(0,0,1,v_K)} \\ \quad + 2 \text{OR}_{(0,0,0,v_K)}, & i = 2, \\ \text{OR}_{(1,1,1,v_K)} - \text{OR}_{(1,1,0,v_K)} \\ \quad - \text{OR}_{(1,0,1,v_K)} - \text{OR}_{(0,1,1,v_K)} \\ \quad + \text{OR}_{(1,0,0,v_K)} + \text{OR}_{(0,1,0,v_K)} \\ \quad + \text{OR}_{(0,0,1,v_K)} - \text{OR}_{(0,0,0,v_K)}, & i = 3, \end{cases}$$

all of them derived by inserting (10) into (13). The first unadjusted excess odds ratio is caused by the joint (marginal and interaction) effect of factors 1, 2 and 3, the second includes second and third order interaction between these three factors but no marginal effects [23, 16], whereas the third only includes the highest third order interaction between factors 1, 2 and 3 [17].

4 Inference of joint effects and interaction

Assume that a case-control dataset $(x_1, Y_1), \dots, (x_n, Y_n)$ of size $n = n_0 + n_1$ is available, with n_0 controls ($Y_a = 0$) and n_1 cases ($Y_a = 1$). Since this is a retrospective sample, we cannot estimate all parameters of the logistic regression model (2). On the other hand, it is possible to estimate the structural parameters ψ consistently when controls are frequency matched with cases. This is the unconditional logistic regression approach, whereby the prospective log likelihood

$$(\hat{\psi}, \hat{\kappa}) = \arg \max_{\psi, \kappa} L(\psi, \kappa) = \arg \max_{\psi, \kappa} \sum_{a=1}^n \log P(Y_a | x_a) \tag{17}$$

is maximized over the model parameters in (2). This approach can also be used if cases and controls are matched in strata, as long as all variables used for matching are included in the model as covariates, and the number of strata does not grow with sample size. On the other hand, conditional logistic regression is more appropriate if the number of strata is large [5].

Let

$$\hat{\xi} = \xi(\hat{\psi}) \quad (18)$$

be the plug-in estimator of (4) obtained from the maximum likelihood estimator (17). In order to produce a confidence interval for ξ with asymptotic coverage probability $1 - \alpha$ we use the delta method in conjunction with some appropriately chosen monotone increasing and differentiable transformation h . This leads to

$$\text{CI} = [h^{-1}(h(\hat{\xi}) - z_{1-\alpha/2}\text{SE}), h^{-1}(h(\hat{\xi}) + z_{1-\alpha/2}\text{SE})], \quad (19)$$

where z_β is the β -quantile of a standard normal distribution and SE is the standard error of $h(\hat{\xi})$. This method relies on asymptotic normality

$$\hat{\psi} \sim \text{AsN}(\psi, \Sigma) \quad (20)$$

of the structural part of (17) for large samples, with $\Sigma = \mathbf{I}(\psi, \kappa)^{-1}$ an asymptotic approximation of the covariance matrix $\text{Cov}(\hat{\psi})$ of $\hat{\psi}$, which equals the inverse of the Fisher information matrix \mathbf{I} [6]. In order to find SE we first approximate the asymptotic variance of $\hat{\xi}$ by $\sigma^2 = \mathbf{D}\Sigma\mathbf{D}^T$, using (20) and the first order Taylor expansion of $\xi(\psi)$, with $\mathbf{D} = \mathbf{D}(\psi) = d\xi(\psi)/d\psi$ and T referring to vector transposition (see Appendix C for an explicit expression of \mathbf{D}). Then one chooses the function h so that the distribution of $h(\hat{\xi})$ is closer to normal than that of $\hat{\xi}$. Typically h is a variance stabilizing transformation that maps the range of ξ to the real line $(-\infty, \infty)$. The Taylor expansion of h gives the asymptotic variance of $h(\hat{\xi})$. By taking the square root of an estimate of this variance, we find the standard error

$$\text{SE} = \sqrt{\widehat{\text{Var}}(h(\hat{\xi}))} = h'(\hat{\xi})\hat{\sigma}$$

of $h(\hat{\xi})$, with $\hat{\sigma}^2 = \hat{\mathbf{D}}\hat{\Sigma}\hat{\mathbf{D}}^T$ an estimate of σ^2 , $\hat{\mathbf{D}} = \mathbf{D}(\hat{\psi})$ an estimate of \mathbf{D} , and $\hat{\Sigma} = \mathbf{I}(\hat{\psi}, \hat{\kappa})^{-1}$ the inverse of the observed Fisher information matrix. In this paper we will use the transformations

$$h(\xi) = \begin{cases} \xi, & \text{if } \xi = \text{EOR}, \\ \log[(1 + \xi)/(1 - \xi)], & \text{if } \xi = \text{AP}, \\ \log(\xi), & \text{if } \xi = \text{SI}, \end{cases} \quad (21)$$

see for instance Rothman [25] and Hössjer et al. [16]. The three functions in (21) map the corresponding indices ξ from their original ranges $((-\infty, \infty)$ for EOR, $(-1, 1)$ for AP, $(0, \infty)$ for SI) to $(-\infty, \infty)$.

5 Analysis of a real dataset

Multiple sclerosis (MS) is a complex and inflammatory disease causing damage to the central nervous system. Its prevalence is over 0.1% in many countries, affecting large regions of the world [21]. There is solid evidence for a genetic component of the disorder, with a major contribution from variants at the human leukocyte antigen (HLA) complex. It is also well known that presence of allele 15 of the HLA-DRB1

Table 1. Number of cases and controls

Study	Cases	Controls
Swedish EIMS study	1308	1858
Swedish GEMS study	3272	2382
Danish study	1474	3469
Norwegian study	211	692
Combined Nordic study	6265	8401

The four Nordic studies are from Hedström et al. [14]

gene is a risk factor, whereas allele 02 of the HLA-A gene has a protective effect [7]. Several studies reveal that environmental factors, in particular smoking, impact the risk of the disease as well [22]. A Swedish study in Hedström et al. [13] demonstrated positive pairwise additive interaction between the two genetic factors, and also between smoking and each genetic factor. These results have more recently been replicated and refined in Hedström et al. [14], by merging case-control studies from several countries. Due to the size of this meta analysis, it was also possible to investigate whether third order interaction was present between the two genetic factors and smoking. In order to illustrate the methodology of this paper, we present some of the findings from the four Nordic studies (see Table 1) of Hedström et al. [14].

Apart from the two genetic factors and smoking, three other covariates (gender, age, study) were also part of the model. This gives a total of 8 covariates, encoded as

$$\begin{aligned}
 x_1 &= \text{HLA-DRB1 (1 for genotypes with a least one copy of,} \\
 &\quad \text{allele 15, 0 otherwise),} \\
 x_2 &= \text{HLA-A (1 for genotypes with no allele 02, 0 otherwise),} \\
 x_3 &= \text{smoking (1=smoker, 0=non-smoker),} \\
 x_4 &= \text{gender (1=female, 0=male),} \\
 x_5 &= \text{age when MS was detected } (\in \{0, 1, \dots, 73\}), \\
 x_6 - x_8 &= \text{study } ((0,0,0)=\text{EIMS}, (1,0,0)=\text{GEMS}, (0,1,0)=\text{Danish}, \\
 &\quad (0,0,1)=\text{Norwegian}).
 \end{aligned}$$

The last three study covariates were only included for the meta analysis, so that

$$p + q = \begin{cases} 5, & \text{for each separate study,} \\ 8, & \text{for the combined Nordic study.} \end{cases}$$

Let

$$\xi = \frac{\text{OR}_{(\mathbf{1}, \mathbf{v}_K)}}{\text{OR}_{(\mathbf{0}, \mathbf{v}_K)}} \quad (22)$$

be the odds ratio for the joint effect of all the risk factors J , when the confounding risk factors in K are fixed at level \mathbf{v}_K , and the remaining covariates are $\{1, \dots, p + q\} \setminus (J \cup K)$. Table 2 lists maximum likelihood estimates (18) of these odds ratios for various choices of risk factors and confounders. The associated confidence intervals (19) are obtained using the logarithmic transformation in (21). Although

Table 2. Point estimates and 95% confidence intervals of the odds ratio (22)

Study	OR for one factor		
	DR15	A2-	sm
EIMS	3.55 (3.05, 4.13)	1.74 (1.50, 2.02)	1.52 (1.30, 1.78)
GEMS	3.70 (3.30, 4.15)	1.79 (1.60, 2.00)	1.62 (1.44, 1.82)
Danish	3.42 (2.99, 3.92)	1.73 (1.51, 1.98)	3.09 (2.70, 3.55)
Norwegian	5.02 (3.50, 7.21)	1.77 (1.24, 2.53)	2.13 (1.50, 3.04)
Combined	3.60 (3.34, 3.87)	1.75 (1.63, 1.88)	2.00 (1.86, 2.15)
Study	OR for two factors		
	DR15, sm	A2-, sm	DR15, A2-
EIMS	5.41 (4.27, 6.85)	2.71 (2.17, 3.40)	6.18 (4.94, 7.75)
GEMS	5.83 (4.93, 6.91)	2.89 (2.45, 3.42)	6.44 (5.46, 7.60)
Danish	10.49 (8.57, 12.84)	5.35 (4.38, 6.52)	5.95 (4.89, 7.23)
Norwegian	10.86 (6.53, 18.07)	3.78 (2.27, 6.27)	8.84 (5.21, 14.99)
Combined	7.11 (6.37, 7.93)	3.51 (3.16, 3.91)	6.28 (5.64, 6.98)
Study	OR for three factors/confounders		
	DR15, A2-lsm	DR15, A2-lsm	DR15, A2-, sm
EIMS	5.62 (4.27, 7.39)	7.72 (5.17, 11.52)	11.23 (7.81, 16.14)
GEMS	7.07 (5.70, 8.77)	5.61 (4.34, 7.27)	9.96 (7.75, 12.79)
Danish	6.57 (5.01, 8.61)	5.27 (3.95, 7.01)	17.95 (13.34, 24.17)
Norwegian	9.02 (4.29, 18.98)	8.66 (4.14, 18.11)	18.27 (8.62, 38.72)
Combined	6.37 (5.55, 7.31)	6.16 (5.20, 7.29)	12.63 (10.73, 14.85)

The sets of risk factors is J and confounders fully controlled for is K . Each column is denoted as J when $K = \emptyset$, and otherwise as $J|K$. The confidence intervals are given in brackets. We use the notation DR15 for presence of allele 15 at HLA-DRB1, A2- for absence of allele 2 at HLA-A, sm for smoker and nsm for non-smoker

MS is a common disorder, its prevalence is small enough to assume that the point estimates and confidence intervals of Table 2 are representative for the relative risks $RR(\mathbf{1}, \mathbf{v}_K, \mathbf{z})/RR(\mathbf{0}, \mathbf{v}_K, \mathbf{z})$ as well.

We find, for instance, that the point estimate of the marginal odds ratio (or relative risk) of having MS in the combined dataset is 3.6 for individuals with the DRB15 risk allele, compared to those that lack this allele. The corresponding marginal odds ratios for absence of the protecting A2 allele and for smoking are 1.75 and 2.0 respectively. Since the joint odds ratios for all pairs of risk factors are much larger than the corresponding marginal odds ratios, there are strong indications of two-way interactions between all pairs of risk factors. There is possibly some three-way interaction between DR15, A2- and smoking as well, since the joint OR for all three factors is higher than the pairwise odds ratios. On the other hand, the OR for the two genetic factors is only higher among smokers than among non-smokers for one study (EIMS).

The estimates of Table 2 motivate further analysis of the MS datasets in Hedström et al [14] based on the attributable proportion, excess odds ratio and synergy index for the three risk factors DR15, A2- and smoking. Table 3 gives confidence intervals for all three quantities when $J = \{1, 2, 3\}$ and $K = \emptyset$, for various choices of i . It confirms a strong joint effect of all three factors, since AP and EOR are both significantly different from 0, and SI is significantly different from 1. For instance, the estimate $\widehat{AP} = 0.92$ for the combined Nordic data set indicates that about 92% of the odds ratio (or disease risk) for smokers with both genetic risk factors, is not present

Table 3. Point estimates and 95% confidence intervals for AP, EOR and SI

Study	AP for DR15, A2-, smoking		
	Joint effects	2nd & 3rd order interaction	3rd order interaction
EIMS	0.91 (0.87, 0.94)	0.59 (0.43, 0.72)	0.39 (0.12, 0.61)
GEMS	0.90 (0.87, 0.92)	0.35 (0.19, 0.50)	-0.02 (-0.29, 0.25)
Danish	0.94 (0.93, 0.96)	0.60 (0.49, 0.70)	0.09 (-0.18, 0.34)
Norwegian	0.95 (0.89, 0.97)	0.66 (0.36, 0.84)	0.22 (-0.33, 0.66)
Combined	0.92 (0.91, 0.93)	0.53 (0.46, 0.60)	0.15 (0.00, 0.29)
Study	EOR for DR15, A2-, smoking		
	Joint effects	2nd & 3rd order interaction	3rd order interaction
EIMS	10.23 (6.15, 14.30)	6.68 (3.05, 10.31)	4.35 (0.33, 8.37)
GEMS	8.96 (6.46, 11.45)	3.50 (1.32, 5.68)	-0.24 (-3.05, 2.58)
Danish	16.95 (11.62, 22.29)	10.85 (6.55, 15.14)	1.55 (-3.41, 6.51)
Norwegian	17.27 (3.55, 30.99)	12.05 (0.82, 23.28)	4.07 (-7.82, 15.95)
Combined	11.63 (9.57, 13.68)	6.74 (5.00, 8.49)	1.88 (-0.16, 3.93)
Study	SI for DR15, A2-, smoking		
	Joint effects	2nd & 3rd order interaction	3rd order interaction
EIMS	undefined	2.88 (1.87, 4.43)	1.74 (1.10, 2.75)
GEMS	undefined	1.64 (1.25, 2.16)	0.97 (0.71, 1.33)
Danish	undefined	2.78 (2.07, 3.72)	1.10 (0.81, 1.49)
Norwegian	undefined	3.31 (1.50, 7.32)	1.31 (0.62, 2.76)
Combined	undefined	2.38 (2.00, 2.84)	1.19 (0.99, 1.44)

The measures either quantify joint marginal and interaction effects ($i = 1$), the total interaction of order 2 and 3 ($i = 2$), or the highest order 3 of interaction ($i = 3$) between the three risk factors DR15, A2- and smoking ($J = \{1, 2, 3\}$). No other covariates are fully controlled for (so that $p = 3$ and $K = \emptyset$). The confidence intervals are given in brackets. Notice that SI is only defined for measures of interaction

among those that lack all three factors. We also find that the total amount of second and third order interaction between DR15, A2- and smoking is strongly significant. In particular, $\widehat{AP} = 0.52$ for the Nordic metapopulation indicates that about half of the disease risk is due to interaction. One the other hand, $\widehat{SI} = 2.38$ tells that the disease risk increment of a smoker with both genetic risk factors, compared to an individual with none of these risk factors, is more than twice the disease risk increment due to marginal effects only. Finally, we find from the rightmost AP, EOR and SI columns of Table 3 that third order interaction between DR15, A2- and smoking is only significant for one dataset (EIMS).

We have also estimated the attributable proportion separately for males and females (data not shown). The results are in quite good agreement with the upper part of Table 3, although the values for males are slightly larger than those for females, and since the gender specific datasets are smaller, the confidence intervals are wider. Such an analysis illustrates how joint effects of and interaction among three factors ($J = \{1, 2, 3\}$) is affected when one controls for a fourth factor ($K = \{4\}$).

6 Discussion

In this paper we studied how a collection J of binary risk factors jointly influence a binary outcome. Our objective was to develop a general framework for quantifying

marginal effects and various orders of interaction between these factors on an additive scale, when stratifying or partially controlling for other confounding variables. This led to the key result; an expansion of the odds ratios as a sum of marginal effects and interaction terms of different orders, with a finite difference and a combinatorial inclusion–exclusion interpretation. We also showed how to use these odds ratio expansions for estimating and producing confidence intervals for the excess odds ratio, attributable proportion and synergy index from a case-control dataset. The methodology was illustrated using a Nordic meta study for multiple sclerosis. The inferential procedure relies on maximum likelihood estimates from a logistic regression model and the delta method. It has been implemented in a Matlab program that is available from the first author upon request.

Our approach makes it possible to stratify for some variables (in K) at various levels, and yet use the whole data set to estimate effect parameters of the other q confounders, for which there is only partial control. But this requires a large data set in order to estimate all $2^p + q = 2^{|J|+|K|} + q$ parameters of the logistic regression model. An alternative and simpler strategy is to use only those observations (x_a, Y_a) for which the variables in K are at a pre-specified level. This gives a smaller model with only $2^{|J|} + q$ parameters to estimate.

The delta based confidence intervals are fast to compute. This is suitable in applications where a number of different putative risk factors are sought for. We have implemented confidence intervals based on resampling as well, using the bias-corrected accelerated percentile method [9, 10, 16]. There is generally good agreement between the resampling and delta-based intervals for odds ratios and attributable proportions, as long as the interaction effects are not too strong, the order of interaction is not too high and the data set is not too small. On the other hand, the delta based confidence intervals tend to be less accurate for excess odds ratios. For this reason we recommend to report excess odds ratios with resampling based confidence intervals. As a topic of future research, it would be of interest to compare the asymptotic accuracy of the delta and resampling based confidence intervals more systematically.

Another object of further study is to develop odds ratio expansions of main effects and interactions when some of the risk factors are continuous [11] and to find analogous expansions for non-binary outcomes.

Acknowledgement

The authors wish to thank the editors and three anonymous reviewers for helpful comments that considerably improved the structure and content of the paper. Ola Hössjer was financially supported by the Swedish Research Council, contract Nr. 621-2013-4633.

A Proof of Theorem 1

Assume without loss of generality that $J = \{1, \dots, p\}$ and $K = \emptyset$. It is possible then to rewrite (9) as

$$\Delta^v \text{OR} = \text{OR}_v - \sum_{\mathbf{0} \leq w < v} \Delta^w \text{OR}. \quad (23)$$

Since $\Delta^{\mathbf{0}} = 1$, and the set $\{0, 1\}^p$ of binary vectors of length p is partially ordered, we can use (9) to compute $\Delta^v \text{OR}$ recursively for all $v \in \{0, 1\}^p$. Such a procedure gives

$$\Delta^v \text{OR} = \sum_{\mathbf{0} \leq \mathbf{w} \leq v} c_{v, \mathbf{w}} \text{OR}_{\mathbf{w}}, \quad (24)$$

for some constants $c_{v, \mathbf{w}}$. In order to verify (7) and (9), it suffices to prove that

$$c_{v, \mathbf{w}} = (-1)^{|\mathbf{v}| - |\mathbf{w}|}, \quad \text{for all } \mathbf{0} \leq \mathbf{w} \leq v. \quad (25)$$

We do this by induction with respect to $v \in \{0, 1\}^p$. Starting with $v = \mathbf{0}$, we notice first that $c_{\mathbf{0}, \mathbf{0}} = 1$, since $\Delta^{\mathbf{0}} \text{OR} = \text{OR}_{\mathbf{0}} = 1$. As a next step, consider a fixed $v > \mathbf{0}$. In order to establish (25) for all \mathbf{w} such that $\mathbf{0} \leq \mathbf{w} \leq v$, we use the induction hypothesis and assume that (25) holds for all (\mathbf{w}, \mathbf{u}) such that $\mathbf{0} \leq \mathbf{w} \leq \mathbf{u} < v$. Then we insert (24)–(25) into each term on the right hand side of (23), and change order of summation. This gives

$$\begin{aligned} \Delta^v \text{OR} &= \text{OR}_v - \sum_{\mathbf{u}: \mathbf{0} \leq \mathbf{u} < v} \Delta^{\mathbf{u}} \text{OR} \\ &= \text{OR}_v - \sum_{\mathbf{u}: \mathbf{0} \leq \mathbf{u} < v} \sum_{\mathbf{w}: \mathbf{0} \leq \mathbf{w} \leq \mathbf{u}} (-1)^{|\mathbf{u}| - |\mathbf{w}|} \text{OR}_{\mathbf{w}} \\ &= \text{OR}_v - \sum_{\mathbf{w}: \mathbf{0} \leq \mathbf{w} < v} \text{OR}_{\mathbf{w}} \sum_{\mathbf{u}: \mathbf{w} \leq \mathbf{u} < v} (-1)^{|\mathbf{u}| - |\mathbf{w}|}. \end{aligned} \quad (26)$$

It is clear from this expansion that $c_{v, v} = 1$, so that (25) holds for $\mathbf{w} = v$. When $\mathbf{w} < v$, we identify the coefficient of $\text{OR}_{\mathbf{w}}$ as

$$\begin{aligned} c_{v, \mathbf{w}} &= - \sum_{\mathbf{u}: \mathbf{w} \leq \mathbf{u} < v} (-1)^{|\mathbf{u}| - |\mathbf{w}|} \\ &= - \sum_{k=0}^{|\mathbf{v}| - |\mathbf{w}| - 1} (-1)^k |\{\mathbf{u}; |\mathbf{u}| = |\mathbf{w}| + k, \mathbf{w} \leq \mathbf{u} \leq v\}| \\ &= - \sum_{k=0}^{|\mathbf{v}| - |\mathbf{w}| - 1} (-1)^k \binom{|\mathbf{v}| - |\mathbf{w}|}{k} \\ &= -[(1-1)^{|\mathbf{v}| - |\mathbf{w}|} - (-1)^{|\mathbf{v}| - |\mathbf{w}|}] \\ &= (-1)^{|\mathbf{v}| - |\mathbf{w}|}, \end{aligned}$$

and this finishes the proof of (25).

B Proof of Proposition 1

The second part of Proposition 1, equation (12), concerns the case when all orders of interaction are included for predicting the odds ratio, i.e. $i = |\mathbf{v}_J|$. This equation follows directly from the fact that the definition of $\text{OR}_{(v_J, v_K), i}$ in (10) is identical to (9) when $i = |\mathbf{v}_J|$.

In order to prove the first part (11) of Proposition 1, when $0 \leq i < |\mathbf{v}_J|$, we insert the definition of $\Delta^{(\mathbf{w}, \mathbf{v}_K)}\text{OR}$ in (7) into equation (10). Then we change the order of summation. This leads to

$$\text{OR}_{(\mathbf{v}_J, \mathbf{v}_K), i} = \sum_{\substack{\mathbf{0} \leq \mathbf{w} \leq \mathbf{v}_J \\ |\mathbf{w}| \leq i}} \text{OR}_{(\mathbf{w}, \mathbf{v}_K)} \sum_{l=0}^{i-|\mathbf{w}|} (-1)^l \binom{|\mathbf{v}_J| - |\mathbf{w}|}{l}. \quad (27)$$

Put $n = |\mathbf{v}_J| - |\mathbf{w}|$ and $m = i - |\mathbf{w}|$, so that $0 \leq m < n$. The combinatorial identity

$$\sum_{l=0}^m (-1)^l \binom{n}{l} = (-1)^m \binom{n-1}{m}$$

can be proved by induction with respect to $m = 0, 1, \dots, n-1$, using Pascal's triangle in each step. It implies that (27) is equivalent to (11), and this completes the proof.

C Expressions for $D(\psi)$

It will be convenient to introduce the notation

$$\begin{aligned} a &= \text{OR}_{(\mathbf{1}, \mathbf{v}_K)}, \\ b &= \text{OR}_{(\mathbf{1}, \mathbf{v}_K), i-1}, \\ c &= \text{OR}_{(\mathbf{0}, \mathbf{v}_K)}, \end{aligned} \quad (28)$$

respectively for the odds ratio when all the risk factors in J are turned on, the prediction of this odds ratio including main effects and interaction terms up to order $i - 1$, and the odds ratio when all the risk factors in J are turned off. The excess odds ratio (14), the attributable proportion (15), the synergy index (16) and the adjusted odds ratio (22) are all different functions

$$\xi = \begin{cases} (a - b)/c, & \xi = \text{EOR}, \\ (a - b)/\max(a, b), & \xi = \text{AP}, \\ (a - c)/(b - c), & \xi = \text{SI}, \\ a/c, & \xi = \text{adjusted OR}, \end{cases} \quad (29)$$

of a, b and c . Hence, in order to find the derivative D of ξ with respect to ψ , for any of these quantities, it suffices to find the derivatives of a, b and c with respect to ψ . To this end, it is convenient to associate with any binary vector \mathbf{u} of length p , another binary vector

$$\mathbf{1}_{\leq \mathbf{u}} = (\mathbf{1}_{\mathbf{w} \leq \mathbf{u}}; \mathbf{0} < \mathbf{w} \leq \mathbf{1}) \quad (30)$$

of length $2^p - 1$. As in the definition of ψ in (3), the indices of $\mathbf{1}_{\leq \mathbf{u}}$ run over all binary vectors $\mathbf{0} < \mathbf{w} \leq \mathbf{1}$ of length p except the zero vector. For a given \mathbf{w} , the corresponding coordinate of $\mathbf{1}_{\leq \mathbf{u}}$ is 1 if $\mathbf{w} \leq \mathbf{u}$ and zero otherwise.

With these preliminaries, we can state the following result, which follows from (28)–(29) and the definition of the odds ratio (5):

Proposition 2. *Let $\xi = \xi(\psi)$ be a measure (4) of joint effects, marginal effects or interaction among the risk factors in J . Assume that it is a function of the three*

quantities a , b and c in (28). The derivative of ξ with respect to the parameter vector $\boldsymbol{\psi}$, is then given by

$$\mathbf{D} = \frac{d\xi(\boldsymbol{\psi})}{d\boldsymbol{\psi}} = \frac{\partial\xi}{\partial a}\mathbf{A} + \frac{\partial\xi}{\partial b}\mathbf{B} + \frac{\partial\xi}{\partial c}\mathbf{C},$$

where

$$\begin{aligned} \mathbf{A} &= \frac{da}{d\boldsymbol{\psi}} = OR_{(1, \mathbf{v}_K)} \mathbf{1}_{\leq(1, \mathbf{v}_K)}, \\ \mathbf{B} &= \frac{db}{d\boldsymbol{\psi}} = \sum_{\substack{0 \leq u \leq 1 \\ |u| \leq i-1}} \sum_{0 \leq w \leq u} (-1)^{|u|-|w|} OR_{(\mathbf{w}, \mathbf{v}_K)} \mathbf{1}_{\leq(\mathbf{w}, \mathbf{v}_K)} \\ &= \sum_{\substack{0 \leq w \leq 1 \\ |w| \leq i-1}} OR_{(\mathbf{w}, \mathbf{v}_K)} \mathbf{1}_{\leq(\mathbf{w}, \mathbf{v}_K)} (-1)^{i-1-|w|} \binom{|J| - 1 - |w|}{i - 1 - |w|}, \\ \mathbf{C} &= \frac{dc}{d\boldsymbol{\psi}} = OR_{(\mathbf{0}, \mathbf{v}_K)} \mathbf{1}_{\leq(\mathbf{0}, \mathbf{v}_K)}. \end{aligned}$$

With a slight abuse of notation, we included binary vectors \mathbf{u} and \mathbf{w} of length $|J|$ (not p) in the definition of \mathbf{B} . In the last step of this equation we interchanged order of summation between \mathbf{u} and \mathbf{w} and then simplified the expression for the inner sum, similarly as in Appendix B.

References

- [1] Agresti, A.: A survey of exact inference for contingency tables. *Stat. Sci.* **7**(1), 131–153 (1992). [MR1173420](#). doi:[10.1214/ss/1177011454](#)
- [2] Agresti, A.: *Categorical Data Analysis*, 3rd edn. Wiley, Hoboken, New Jersey (2013). [MR3087436](#)
- [3] Andersson, T., Alfredsson, L., Källberg, H., Zdravkovic, S., Ahlbom, A.: Calculating measures of biological interaction. *Eur. J. Epidemiol.* **20**, 575–579 (2005). doi:[10.1007/s10654-005-7835-x](#)
- [4] Assman, S.F., Hosmer, D.W., Lemeshow, S., Mundt, K.A.: Confidence intervals for measures of interaction. *Epidemiology* **7**(3), 286–290 (1996). doi:[10.1097/00001648-199605000-00012](#)
- [5] Breslow, N.E., Day, N.E.: In: *Statistical Methods in Cancer Research. Volume 1. The Analysis of Case-Control Studies*, Lyon, France (1980)
- [6] Casella, G., Berger, R.L.: In: *Statistical Inference*, Thompson Learning (2002)
- [7] Cree, B.: Multiple sclerosis genetics. In: *Handb. Clin. Neurol.*, vol. 122, pp. 193–209 (2014). doi:[10.1016/B978-0-444-52001-2.00009-1](#)
- [8] Ebers, G.C., Sadovnick, A.D.: The geographic distribution of multiple sclerosis: a review. *Neuroepidemiology* **1993** (2013)
- [9] Efron, B.: Better bootstrap confidence intervals (with discussion). *J. Am. Stat. Assoc.* **82**, 171–200 (1987). doi:[10.1080/01621459.1987.10478410](#)
- [10] Efron, B., Tibshirani, R.J.: *An Introduction to the Bootstrap*. Monogr. Stat. Appl. Probab., vol. 57. Chapman and Hall/CRC (1994). [MR1270903](#). doi:[10.1007/978-1-4899-4541-9](#)
- [11] Eide, G.E.: Attributable fractions: fundamental concepts and their visualization. *Stat. Methods Med. Res.* **10**, 159–193 (2001). doi:[10.1177/096228020101000302](#)

- [12] Greenland, S.: Additive risk versus additive relative risk models. *Epidemiology* **4**(1), 32–36 (1993). doi:[10.1097/00001648-199301000-00007](https://doi.org/10.1097/00001648-199301000-00007)
- [13] Hedström, A.K., Sundqvist, E., Bäärnhielm, M., Nordin, N., Hiller, J., Kockum, I., Olsson, T., Alfredsson, L.: Smoking and two human leukocyte antigen genes interact to increase the risk for multiple sclerosis. *Brain* **134**, 653–664 (2011). doi:[10.1093/brain/awq371](https://doi.org/10.1093/brain/awq371)
- [14] Hedström, A.K., Katsoulis, M., Hössjer, O., Bomfim, I.L., Oturai, A., et al.: The interaction between smoking and HLA genes in multiple sclerosis; replication and refinement. *Eur. J. Epidemiol.* (2017), accepted for publication
- [15] Hosmer, D.W., Lemeshow, S.: Confidence interval estimation of interaction. *Epidemiology* **3**(5), 452–456 (1992). doi:[10.1097/00001648-199209000-00012](https://doi.org/10.1097/00001648-199209000-00012)
- [16] Hössjer, O., Kockum, I., Alfredsson, L., Hedström, A-K., Olsson, T., Lekman, M.: A new normalization and generalized definition of attributable proportion. *Epidemiol. Methods* (2017) (available online December 2016). doi:[10.1515/em-2015-0028](https://doi.org/10.1515/em-2015-0028)
- [17] Katsoulis, M., Bamia, C.: Moving from two- to multi-way interactions between binary risk factors on the additive scale. Submitted for publication (2017)
- [18] Knol, M.J., VanderWeele, T.J., Groenwold, R.H., Klungel, O.H., Rovers, M.M., Grobbee, D.E.: Estimating measures of interaction on an additive scale for preventive exposures. *Eur. J. Epidemiol.* **26**, 433–438 (2011). doi:[10.1007/s10654-011-9554-9](https://doi.org/10.1007/s10654-011-9554-9)
- [19] Lekman, M., Hössjer, O., Andrew, P., Källberg, H., Uvehag, D., Charney, D., Manji, H., Rush, J., McMahon, F., Moore, J., Kockum, I.: The genetic interacting landscape of 63 candidate genes in Major Depressive Disorder: an explorative study *Biodata Min.* **7**, 19 (2014) (18 pages)
- [20] McCullagh, P., Nelder, J.: *Generalized Linear Models*, 2nd edn. Chapman & Hall, London (1989)
- [21] Multiple Sclerosis International Federation: *Atlas of MS Survey* (2013)
- [22] Olsson, T., Barcellos, L.F., Alfredsson, L.: Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nat. Rev. Neurol.* **13**(1), 25–36 (2017). doi:[10.1038/nrneurol.2016.187](https://doi.org/10.1038/nrneurol.2016.187)
- [23] Petti, S., Masood, M., Scully, C.: The magnitude of tobacco smoking-betel quid chewing drinking interaction effect on oral cancer in South-East Asia. A meta-analysis of observational studies *PLoS ONE* **8**(11), e78999 (2013)
- [24] Prentice, R.L., Pyke, R.: Logistic disease incidence models and case-control studies. *Biometrika* **66**(3), 403–411 (1979). doi:[10.1093/biomet/66.3.403](https://doi.org/10.1093/biomet/66.3.403)
- [25] Rothman, K.J.: The estimation of synergy or antagonism. *Am. J. Epidemiol.* **103**, 506–511 (1976). doi:[10.1093/oxfordjournals.aje.a112252](https://doi.org/10.1093/oxfordjournals.aje.a112252)
- [26] Rothman, K.J.: *Epidemiology. An Introduction*, 2nd edn. Oxford Univ. Press, N.Y. (2012)
- [27] Rothman, K.J., Greenland, S., Walker, A.M.: Concepts of interaction. *Am. J. Epidemiol.* **112**, 467–470 (1980). doi:[10.1093/oxfordjournals.aje.a113015](https://doi.org/10.1093/oxfordjournals.aje.a113015)
- [28] Skrdal, A.: Interaction as departure from additivity in case-control studies: a cautionary note. *Am. J. Epidemiol.* **158**, 251–258 (2003). doi:[10.1093/aje/kwg113](https://doi.org/10.1093/aje/kwg113)
- [29] VanderWeele, T.J.: Reconsidering the denominator of the attributable proportion for interaction. *Eur. J. Epidemiol.* **28**, 779–784 (2013). doi:[10.1007/s10654-013-9843-6](https://doi.org/10.1007/s10654-013-9843-6)
- [30] VanderWeele, T.J., Robins, J.M.: Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika* **95**, 49–61 (2008). doi:[10.1093/biomet/asm090](https://doi.org/10.1093/biomet/asm090)