On the Existence of Suitable Models for Additive Interaction with Continuous Exposures

by

Daniel Berglund, Claudia Carlucci, Helga Westerlind,
and Timo Koski

Abstract

Additive interaction can be of importance for public health interventions and it is commonly defined using binary exposures. There has been expansions of the models to also include continuous exposures, which could lead to better and more precise estimations of the effect of interventions. In this paper we define the intervention for a continuous exposure as a monotonic function. Based on this function for the interventions we prove that there is no model for estimating additive interactions with continuous exposures for which it holds that; (i) both exposures have marginal effects and no additive interaction on the exposure level for both exposures, (ii) neither exposure has marginal effect and there is additive interaction between the exposures. We also show that a logistic regression model for continuous exposures will always produce additive interaction if both exposures have marginal effects.

Keywords: Additive Interaction; Multiplicative Interaction; Logistic Regression; Linear Odds; Continuous Exposures; Public Health; Interventions

1 Introduction

For some diseases and traits, a combination of factors is required for the diseases to occur, and factors can also modify each others effect, and significantly increase or decrease in strength. In epidemiological and social science research much work has been focused on finding and estimating such effects [20, 28, 32]. For example smoking combined with the genotype HLA-DRB1 SE substantially increases the risk for Rheumatoid Arthritis [11]. However, the exact mechanisms and potential mechanistic interactions are in general hard to define and impossible to estimate from data. This is referred to as biological, or sufficient

cause, interaction among others, and does not necessarily imply that some mechanisms behind the factors are directly interacting, such as a chemical reaction [20, 28].

Statistical interaction on the other hand has a clear definition, it is when the effect from one exposure depends on the level of another exposure [20, 28]. Statistical interactions are scale dependent and different scales can lead to different conclusions. An example is the interaction between smoking and asbestos on the risk of lung cancer where the risk from asbestos is higher for nonsmokers than for smokers on the ratio scale, however the opposite is true on the additive scale [20].

For public health it can be argued that the additive scale is the preferred scale as it leads to the most effective interventions when choosing which group to intervene upon [1, 4, 20, 21, 23, 33]. The additive scale is also connected to the sufficient-cause model, where the presence of interaction on the additive scale implies interaction in the sufficient-cause model, and thereby that both factors are involved in the causal pathway [28, 30, 32–34].

Most research on additive interaction have focused on binary exposures with few papers in the literature related to additive interaction using continuous exposures [3, 12, 13, 29, 31]. Instead of dichotomizing the continuous exposures, as in [3, 29], in the ideal case the ideal model would use the continuous exposures directly. Such an approach could improve the models and lead to new insights, as dichotomizing can cause a loss of information [22].

We will show in this paper that there is no model for additive interaction with continuous exposures that met the criteria for a suitable model. Our result also implies that the models with continuous exposures for additive interaction in [12, 13, 31] are flawed.

In Section 2 we summarize some of the background for interaction when the exposures are binary. Then in Section 3 we will define interventions and interaction for continuous exposures. In that section we also show the main result; that there is no model that can meet the criteria for a suitable model.

2 Interaction with Binary Exposures

We are going to start with a summary of the background for additive and multiplicative interaction with binary exposures. Let D denote a binary outcome, e.g., disease, and let X and Y be two binary exposures. We assume that there is adjustment for confounding, and that after the adjustment there is no confounding, this is know as the conditional exchangeability assumption [28].

Using the potential outcomes model [8, 25], and with exposure levels set to $X = x$ and $Y = y$ let

$$p_{x,y} = P(D = 1|X = x, Y = y). \tag{2.1}$$

We make two additional assumptions; First, for the causal effects to be identifiable we require the positivity assumption to be true, i.e., that no $p_{x,y}$ is either zero or one [17, 28]. Second, the effect of the exposures needs to be consistent for all individuals [18, 28];

Formally, if we would intervene upon an exposure for an individual the effect on D would be the same as if the individual would have had that exposure level in the first place [18].

To make it easier to interpret and to be able to estimate the additive interaction in case-control studies relative risks (RR) and odds ratios (OR) are used instead of the probabilities. The relative risks can be approximated with odds ratios either if case-control data is used and the rare disease assumption is met or if certain study designs are used [14]. Note that the rare disease assumption applies to the risk in all stratas studied, and not only the prevalence in the general population as a rare disease in the general population might not be rare in some specific strata.

The relative risks and odds ratios for the exposures $X = x$ and $Y = y$ with the reference group *ref* are

$$\text{RR}_{x,y}^{ref} = \frac{p_{x,y}}{p_{ref}} \tag{2.2}$$

and

$$\text{OR}_{x,y}^{ref} = \frac{\frac{p_{x,y}}{1-p_{x,y}}}{\frac{p_{ref}}{1-p_{ref}}}. \tag{2.3}$$

We now define what interventions on binary exposures are based on previous literature [20].

Definition 2.1: *Intervention on binary exposure*
We define the intervention on the binary exposures X and Y as changing the population's exposures with some binary values x_d and y_d respectively from the reference level x_0, y_0. Then

$$p_{x_0+x_d,y_0+y_d} - p_{x_0,y_0} \tag{2.4}$$

is the effect of the intervention. ■

As mentioned in the Introduction, interactions are scale dependent, and the two most common scales are multiplicative (ratio scale) and additive. As the name suggests, in the multiplicative model for interaction there is interaction if the effect for the doubly exposed group does not follow multiplicative scaling, while in the additive model there is interaction if the effect does not follow additive scaling. The main criticism against multiplicative interaction is that it is not connected to any actual interaction on the biological level [20, 31]. However one could make similar arguments against additive interaction, as even if additive interaction implies sufficient cause interaction, sufficient cause interaction does not imply any chemical or biological interaction [20]. Additive interaction can occur without interaction, for instance from competition between exposures, e.g., an individual can not die from cancer if they died in a car crash [6]. It is recommended that both additive and multiplicative measures are estimated in studies [15, 31].

2.1 Multiplicative Interaction

The multiplicative interaction measure between two exposures is defined as the following [20, 31].

Definition 2.2: *Multiplicative interaction*
The multiplicative interaction measure for two exposures is

$$\mathrm{M} = \frac{p_{11}p_{00}}{p_{10}p_{01}}. \tag{2.5}$$

■

If the measure is one then there is no multiplicative interaction. Expressed using relative risks, multiplicative interaction is

$$\mathrm{M} = \frac{\mathrm{RR}_{11}}{\mathrm{RR}_{10}\mathrm{RR}_{01}}. \tag{2.6}$$

The reference group for the relative risks are cancelled out so the choice of reference does not matter, and the value of the measure also is the same as when probabilities are used.

2.2 Additive Interaction

Additive interaction can be derived based on interventions [20]. If we would intervene on Y by setting $Y = 1$, for either the individuals with $X = 0$ or with $X = 1$, in which group would the intervention have the most effect? The effect for $X = 0$ is $p_{01} - p_{00}$ and the effect for $X = 1$ is $p_{11} - p_{10}$; thus the difference between these two effects is $(p_{01} - p_{00}) - (p_{11} - p_{10})$. Rearranging leads to the interaction contrast (IC),

$$\mathrm{IC} = p_{11} - p_{10} - p_{01} + p_{00}. \tag{2.7}$$

IC measures the amount of additive interaction: if there is no difference between the effects (i.e., no additive interaction) then the contrast is zero. If there is interaction on the additive scale then IC represents the effect that is lost (or gained) by intervening on the wrong (or correct group). Negative IC is referred to as subadditive interaction while positive IC is referred to as superadditive interaction [30]. An interesting note is that it does not matter whether the intervention is considered on X or Y, or if the intervention is to set the exposure to zero or one, the value of IC is the same, only its sign changes.

Additive interaction can also be derived from the sufficient cause model and different inequalities using IC can also imply various types of sufficient cause interaction, such as synergism between X and $\bar{Y}$ [20, 28, 32]. $IC \neq 0$ implies the presence of interaction, but is not a necessary condition, as there can be interaction in the true underlying sufficient cause model even if $\mathrm{IC} = 0$ [20].

Using relative risks for additive interaction the following measures of additive interaction can be derived from the IC [20].

$$\mathrm{RERI} = \mathrm{RR}_{11}^{00} - \mathrm{RR}_{10}^{00} - \mathrm{RR}_{01}^{00} + 1 \tag{2.8}$$

$$\text{AP} = \frac{\text{RERI}}{\text{RR}_{11}^{00}} \tag{2.9}$$

$$\text{AP}^* = \frac{\text{RERI}}{\text{RR}_{11}^{00} - 1} \tag{2.10}$$

The interpretation of the relative excess risk due to interaction (RERI) is similar to that of IC. RERI measures how much more, or less, risk there is in the doubly exposed group compared to if there was no additive interaction [20]. The attributable proportion(AP) is the proportion of the *disease* in the group with both exposures that is due to the interaction, while AP* is the proportion of the *effect* that is due to interaction [27]. In other words the two AP measures tells us what proportion of the outcome is caused by the interaction. Similar measures can also be calculated for the exposures themselves, e.g., how much of the disease is caused by exposure to x alone [10, 27].

The measures are defined using p_{00} as the reference group and the exposures as harmful. In the case any exposure is protective the reference group needs to be switched so that all the relative risks are risk ratios, otherwise the interaction measure can be incorrect and misleading [16].

2.3 Modeling Relative Risks and Odds Ratios

There are several possible ways to model the probabilities, relative risks or odds ratios. Following [5] with some changes and some abuse of notation we define two possible forms for the models. With $r(x)$ as a risk, rate, odds or similar, the additive form and multiplicative form of the model are

$$r(x) = \alpha + x\beta \tag{2.11}$$

and

$$r(x) = \alpha(1 + x\beta). \tag{2.12}$$

The difference between the models is the effect of the parameters; if $r(x)$ changes in an additive way (first equation), or multiplicative way (second equation). Note that the additive form of the model in Equation 2.11 is the one implied by "no causal interaction" in Rothman's model for causality [5, 20]. Under certain conditions, depending on the covariates, models of the multiplicative form can be used to make inferences even if the true model's form is additive [5].

Logistic Regression

Odds ratios are often modeled and adjusted for confounding by using logistic regression with the model shown below [2, 9, 20, 31].

$$\ln\left(\frac{p_{x,y}}{1 - p_{x,y}}\right) = \alpha + \beta_1 x + \beta_2 y + \beta_3 xy \tag{2.13}$$

The logistic regression model is a multiplicative model since the logarithmic transformation means that the coefficients have a multiplicative effect on the odds. With the logistic model the corresponding measures, with the unexposed group as reference, for additive interaction are:

$$\text{RERI}_{\text{LR}} = e^{\beta_1+\beta_2+\beta_3} - e^{\beta_1} - e^{\beta_2} + 1 \tag{2.14}$$

$$\text{AP}_{\text{LR}} = \frac{e^{\beta_1+\beta_2+\beta_3} - e^{\beta_1} - e^{\beta_2} + 1}{e^{\beta_1+\beta_2+\beta_3}} \tag{2.15}$$

$$\text{AP}^*_{\text{LR}} = \frac{e^{\beta_1+\beta_2+\beta_3} - e^{\beta_1} - e^{\beta_2} + 1}{e^{\beta_1+\beta_2+\beta_3} - 1} \tag{2.16}$$

Testing for the presence of additive interaction using logistic regression can not be directly done by testing the interaction coefficient, β_3, since $\beta_3 = 0$ does not have to imply that interaction on the additive scale is not present [9, 24]. Instead other methods such as the delta method [9], bootstrap methods [2] or MOVER [39] can be used. It is worth noting that methods using the delta method have shown worse accuracy for the confidence interval in simulations than other methods [2, 39]. Also, the logistic model can have problems when including covariates without also including the interaction terms between the exposures and the covariates. If this is the case, it implies there is additive interaction between the exposures and the covariates [7, 24].

Linear Odds

The linear odds model, Equation 2.17, is an alternative to logistic regression, which is additive instead of multiplicative. The linear odds model uses the odds instead of the log odds like the logistic model,

$$\frac{p_{x,y}}{1-p_{x,y}} = a + b_1 x + b_2 y + b_3 xy. \tag{2.17}$$

The linear odds model is useful for assessing additive interaction since when used together with AP^* the additive interaction measures does not have the problems of misspecification and uniqueness [24]. However, the downside is that this model can lead to negative odds and fail to converge [24, 37, 38]. Furthermore, the maximum likelihood estimators can have problems when used with continuous covariates [37].

With the linear odds model, and the unexposed group as reference, the additive interaction measures are:

$$\text{RERI}_{\text{LO}} = \frac{b_3}{a} \tag{2.18}$$

$$\text{AP}_{\text{LO}} = \frac{b_3}{a + b_1 + b_2 + b_3} \tag{2.19}$$

$$\text{AP}^*_{\text{LO}} = \frac{b_3}{b_1 + b_2 + b_3} \tag{2.20}$$

The presence of additive interaction in the measures only depend on b_3. Testing $b_3 \neq 0$ is simple since the coefficients are normally distributed and the test can be done with a Wald test or using bootstrap methods [19, 24, 36].

3 Interaction with Continuous Exposures

For this section we will be using the continuous exposures G and E instead of the binary exposures X and Y as in the previous section. As we will see multiplicative interaction is largely the same as in the binary case. However, we will show later in the section 3.2 that for additive interaction there is no suitable model for the probabilities when using continuous exposures.

3.1 Multiplicative Interaction

The extension of multiplicative interaction to also cover continuous variables is straightforward. We are going to use logistic regression, since the multiplicative interaction is independent of the covariates. Adapting Equation 2.6 with the logistic regression model with the continuous exposures leads to

$$\begin{aligned}\mathrm{M}_{g,e} =& \frac{e^{\beta_1(g_1-g_0)+\beta_2(e_1-e_0)+\beta_3(g_1e_1-g_0e_0)}}{e^{b_1(g_1-g_0)+b_3e_0(g_1-g_0)}e^{b_2(e_1-e_0)+b_3g_0(e_1-e_0)}} \\ =& \frac{e^{\beta_3(g_1e_1-g_0e_0)}}{e^{\beta_3(e_0(g_1-g_0)+g_0(e_1-e_0))}} \\ =& e^{\beta_3(g_1-g_0)(e_1-e_0)}.\end{aligned} \tag{3.1}$$

$\mathrm{M}_{g,e}$ only depends on β_3 assuming that $g_1 - g_0 \neq 0$ and $e_1 - e_0 \neq 0$ $M_{g,e}$. Therefore, the presence of interaction and its direction can be tested with a Wald test or bootstrap methods. Assuming that (g, e) is not random then because $\hat{\beta} \sim N(\mu_\beta, \Sigma_\beta)$ the confidence interval with confidence level α for $ln(M_{g,e})$ is:

$$\mathrm{CI}_{\ln(\mathrm{M}_{g,e})} = (g_1 - g_0)(e_1 - e_0)(\hat{\beta}_3 \pm \lambda_{\alpha/2}\hat{s}_{\beta_3}), \tag{3.2}$$

where the parameters, $\hat{\beta}_3$, $\hat{s}_{\beta_3}$ are estimated from the data, and $\lambda_{\alpha/2}$ is the normal quantile for the confidence level α. Since the exponential function is strictly increasing it follows that

$$\mathrm{CI}_{\mathrm{M}_{g,e}} = e^{(g_1-g_0)(e_1-e_0)(\hat{\beta}_3 \pm \lambda_{\alpha/2}\hat{s}_{\beta_3})}. \tag{3.3}$$

The confidence interval is the smallest when the point g_1, e_1 is close to the reference point and increases exponentially as it moves away from the reference point.

3.2 Additive Interaction

Additive interaction for continuous exposures can be derived in a similar manner as for the binary case. However, the connection with the sufficient cause model no longer holds, since the sufficient cause model is undefined for continuous exposures [32].

We start by defining the interventions on the exposures, which now have to be more general instead of just setting the exposure to 1 or 0.

Definition 3.1: *Intervention on continuous exposures*
The intervention on the continuous exposures G and E are functions, $I_G(g)$ and $I_E(e)$, that transforms the exposures for all individuals in the population. Let g_0, e_0 be the individual's exposure levels. The effect of the intervention is then defined as

$$p_{I_G(g_0),I_E(e_0)} - p_{g_0,e_0}. \tag{3.4}$$

■

If G and E are binary then this would be equivalent to the binary definition since the only two possible functions are to either remove the exposure, or expose everyone. In the continuous case, a possible function for $I_G(g)$ and $I_E(e)$ could for instance be that the intervention on $X = x_0$ for some constant x_d is $I_X(x_0) = x_0 - x_d$ as used in [12, 13, 31]. In this paper we will unless otherwise noted, set the interventions as some strictly decreasing function.

Before defining additive interaction for continuous exposures we will also define the marginal effect of an intervention.

Definition 3.2: *Marginal effect*
The marginal effect from an intervention, $I_G(g_0)$, on a continuous exposure G is defined as

$$p_{I_G(g_0)} - p_{g_0}. \tag{3.5}$$

An intervention, $I_G(g)$, has no marginal effect if

$$\forall g_0 \in \mathbb{R}\ p_{I_G(g_0)} - p_{g_0} = 0. \tag{3.6}$$

■

Just as in the binary case we can derive the IC for the continuous case as the effect difference between intervening on G on either the groups defined by $E = e_A$ or $E = e_B$, or intervening on E on the groups defined by $G = g_A$ or $G = g_B$. Formally

$$IC_{I_G(g_0),e_A,g_0,e_B} = (p_{I_G(g_0),e_A} - p_{g_0,e_A}) - (p_{I_G(g_0),e_B} - p_{g_0,e_B}) \tag{3.7}$$

and

$$IC_{g_A,I_E(e_0),g_B,e_0} = (p_{g_A,I_E(e_0)} - p_{g_A,e_0}) - (p_{g_B,I_E(e_0)} - p_{g_B,e_0}). \tag{3.8}$$

However, in contrast to the binary case the additive interaction measure is no longer unique, since instead of choosing between intervening on two different groups, based on

a binary exposure, the scale is now continuous and there are multiple ways to define the groups. Hence, there are several possible types of additive interaction.

We will start with the simplest case; the case when there is no additive interaction for some *single* value of the exposure intervened upon. In other words, for that value of the exposure it does not matter for the intervention what value to the other exposure takes.

Definition 3.3: *Single additive interaction*
There is single additive interaction for $G = g_0$ with the intervention $I_G(g)$ if

$$\exists e_A, e_B \in \mathbb{R}\ IC_{I_G(g_0),e_A,g_0,e_B} \neq 0, \tag{3.9}$$

and for $E = e_0$ with the intervention $I_E(e)$,

$$\exists g_A, g_B \in \mathbb{R}\ IC_{g_A,I_E(e_0),g_B,e_0} \neq 0. \tag{3.10}$$

With no single additive interaction being the opposite,

$$\forall e_A, e_B \in \mathbb{R}\ IC_{I_G(g_0),e_A,g_0,e_B} = 0, \tag{3.11}$$

and

$$\forall g_A, g_B \in \mathbb{R}\ IC_{g_A,I_E(e_0),g_B,e_0} = 0. \tag{3.12}$$

■

If there, for all values of the exposure is no single additive interaction we say that there is no exposure additive interaction for that particular exposure.

Definition 3.4: *Exposure additive interaction*
There is exposure additive interaction for G with the intervention $I_G(g)$ if

$$\exists g_0, e_A, e_B \in \mathbb{R}\ IC_{I_G(g_0),e_A,g_0,e_B} \neq 0, \tag{3.13}$$

and for E with the intervention $I_E(e)$,

$$\exists g_A, g_B, e_0 \in \mathbb{R}\ IC_{g_A,I_E(e_0),g_B,e_0} \neq 0. \tag{3.14}$$

So there is no exposure additive interaction for G if

$$\forall g_0, e_A, e_B \in \mathbb{R}\ IC_{I_G(g_0),e_A,g_0,e_B} = 0, \tag{3.15}$$

and for E if

$$\forall g_A, g_B, e_0 \in \mathbb{R}\ IC_{g_A,I_E(e_0),g_B,e_0} = 0. \tag{3.16}$$

■

For the binary model both definitions correspond to absence of additive interaction. Exposure additive interaction is what we would consider closest to being equivalent to the intuitive idea of additive interaction with binary exposures, namely that the effect of the intervention on an exposure does not depend on the level of another exposure no matter what the levels of the first and second exposures are.

Existence of a Suitable Model for the Probabilities

So far with the continuous exposures we have not defined any model for the probabilities, $p_{g,e}$. In this section we are going to show the main result of this paper, Theorem 3.4, which shows that there is no model for the probabilities that can fulfill the criteria we consider required for the model.

We propose the following two criteria that the model should be able to fulfill:

i. There exist some values for the model parameters for which both exposures have marginal effects and no additive interaction on the exposure level for both exposures.

ii. There exist some values for the model parameters for which there are no marginal effects and there is additive interaction between the exposures.

which, formally based on the definitions of marginal effect and exposure additive interaction, and with interventions $I_G(g_0)$, $I_E(e_0)$, is given in the following definition.

Definition 3.5: *Suitable model*
A suitable model meet both of the following criteria:

i. There exist some values for the model parameters, K_1, for which all of the following statements hold:

$$\exists g_0 \in \mathbb{R}\ p_{I_G(g_0)} - p_{g_0} \neq 0 \tag{3.17}$$

$$\exists e_0 \in \mathbb{R}\ p_{I_E(e_0)} - p_{e_0} \neq 0 \tag{3.18}$$

$$\forall g_0, e_A, e_B \in \mathbb{R}\ IC_{I_G(g_0),e_A,g_0,e_B} = 0 \tag{3.19}$$

$$\forall g_A, g_B, e_0 \in \mathbb{R}\ IC_{g_A,I_E(e_0),g_B,e_0} = 0 \tag{3.20}$$

ii. There exist some values for the model parameters, K_2, for which all of the following statements hold:

$$\forall g_0 \in \mathbb{R}\ p_{I_G(g_0)} - p_{g_0} = 0 \tag{3.21}$$

$$\forall e_0 \in \mathbb{R}\ p_{I_E(e_0)} - p_{e_0} = 0 \tag{3.22}$$

$$\exists g_0, e_A, e_B \in \mathbb{R}\ IC_{I_G(g_0),e_A,g_0,e_B} \neq 0 \tag{3.23}$$

$$\exists g_A, g_B, e_0 \in \mathbb{R}\ IC_{g_A,I_E(e_0),g_B,e_0} \neq 0 \tag{3.24}$$

■

Note that the model need to fulfill both criteria to be considered as suitable, but not with the same values for the model parameters. E.g., if for some model k_i is the parameter for the interaction, then criterion (i) means that $k_i = 0$, while criterion (ii) implies that $k_i \neq 0$.

We will now show that with the intervention functions as strictly decreasing there is no model for the probabilities such that both criteria can be fulfilled. To prove this result in Theorem 3.4 we first require some additional lemmas. To start with we are going to derive alternative conditions for the conditions in Definition 3.3 and 3.4.

Lemma 3.1: *There is no single additive interaction for $I_G(g_o)$ if*

$$\forall e \in \mathbb{R}: \ \frac{\partial p(I_G(g_0), e)}{\partial e} = \frac{\partial p(g_0, e)}{\partial e}, \tag{3.25}$$

and correspondingly for $I_E(e_o)$,

$$\forall g \in \mathbb{R}: \ \frac{\partial p(g, I_E(e_0))}{\partial g} = \frac{\partial p(g, e_0)}{\partial g}. \tag{3.26}$$

Proof: Proving the case with intervention on G. The condition $IC_{I_G(g_0),e_A,g_0,e_B} = 0$ can be written as

$$p_{I_G(g_0),e_A} - p_{g_0,e_A} = p_{I_G(g_0),e_B} - p_{g_0,e_B}, \tag{3.27}$$

which means that since the condition in Definition 3.3 is *for all* e_A, e_B for some constant α we can instead write the condition as

$$\forall e \in \mathbb{R} \ p_{I_G(g_0),e} - p_{g_0,e} = \alpha. \tag{3.28}$$

Then since $p_{I_G(g_0),e} - p_{g_0,e}$ is a constant function of e its derivative has to be

$$\forall e \in \mathbb{R} \ \frac{\partial}{\partial e}(p_{I_G(g_0),e} - p_{g_0,e}) = 0, \tag{3.29}$$

which leads to

$$\forall e \in \mathbb{R}: \ \frac{\partial p(I_G(g_0), e)}{\partial e} = \frac{\partial p(g_0, e)}{\partial e} \tag{3.30}$$

as a condition equivalent to Equation 3.11. ■

Lemma 3.2: *There is no exposure additive interaction for G if*

$$\forall e, g_0 \in \mathbb{R}: \ \frac{\partial p(I_G(g_0), e)}{\partial e} = \frac{\partial p(g_0, e)}{\partial e}, \tag{3.31}$$

and correspondingly for E,

$$\forall g, e_0 \in \mathbb{R}: \ \frac{\partial p(g, I_E(e_0))}{\partial g} = \frac{\partial p(g, e_0)}{\partial g}. \tag{3.32}$$

Proof: The result follows directly from using the the definition of exposure additive interaction with the proof in Lemma 3.1. ■

Based on the above lemmas we can derive the form which the model for $p_{g,e}$ must have to be able to show no exposure additive interaction.

Theorem 3.3: *Let k_e and m be constants, h some function of G, and $f(g, e)$ is the interaction function and k_i is the parameter for the interaction. For the model to be able to show no exposure additive interaction for an exposure $G = g$ relative to some other exposure $E = e$ it is required that the model can be written on the form*

$$p(g, e) = m + k_e e + h(g) + k_i f(g, e). \tag{3.33}$$

If and only if $k_i = 0$ then there is no additive interaction on the exposure level.

Proof: For Equation 3.31 in the proof of Lemma 3.2 to hold it must be true that for a constant, k_e,

$$\frac{\partial p(g, e)}{\partial e} = k_e. \tag{3.34}$$

Then with constants k_e, m and function $h(g)$, the model for $p(g, e)$ can be written on the form of

$$p(g, e) = m + k_e e + h(g). \tag{3.35}$$

More generally and since we want to model the interaction there could be some parameter for an interaction function that is zero. I.e., with k_i as the parameter for the function f which represents the interaction, $p(g, e)$ can be written as

$$p(g, e) = m + k_e e + h(g) + k_i f(g, e). \tag{3.36}$$

From Lemma 3.2 we know that for there to be no additive interaction on the exposure level, then k_i must be 0. ■

From these results it also follows that for two exposures with both being able to show no exposure additive interaction the model's form has to be

$$p(g, e) = m + k_g g + k_e e + k_i f(g, e). \tag{3.37}$$

Using the form of the model above that is required for the model to be able to meet criterion (i) we will now show that the model can not also fulfill criterion (ii).

Theorem 3.4: *Let the intervention functions be strictly decreasing, then there is no model which can fulfill both of the criteria in Definition 3.5.*

Proof: We prove the theorem by showing that the combination of critera (i) and (ii) leads to a contradiction. Assume that the model meet both criteria (i) and (ii). From the first criterion we know from Theorem 3.3 that the model has to have the form of Equation 3.37. From Definition 3.1 the marginal effects for g and e are,

$$p_{I_G(g_0),e} - p_{g_0,e} = k_g(I_G(g_0) - g_0) + k_i(f(I_G(g_0), e) - f(g_0, e)), \tag{3.38}$$

and

$$p_{g,I_E(e_0)} - p_{g,e_0} = k_e(I_E(e_0) - e_0) + k_i(f(g, I_E(e_0)) - f(g, e_0)). \tag{3.39}$$

No marginal effects for an exposure means that the effect from the intervention on that exposure is zero,

$$\forall g_0, e: \; p_{I_G(g_0),e} - p_{g_0,e} = 0, \tag{3.40}$$

and

$$\forall g, e_0: \; p_{g,I_E(e_0)} - p_{g,e_0} = 0. \tag{3.41}$$

We will be focusing on the exposure G in the next part, but the result also holds for E. Criterion (ii) requires that there is interaction, formally, $k_i \neq 0$ and $\exists g, e : f(g,e) \neq 0$. Then Equation 3.38 have to depend on the interaction function, unless it holds that

$$\forall g_0, e: \; f(I_G(g_0), e) - f(g_0, e) = 0. \tag{3.42}$$

If Equation 3.42 is true then it implies that the value of g does not matter for $f(g,e)$ for some given e. The function $f(g,e)$ can then not represent the interaction, since the definition of statistical interaction requires that the effect from one exposure depends on the level of another exposure. This in turn implies that $p(g,e)$ has no interaction. However, this is a contradiction since criterion (ii) states that there is interaction. ■

Model Example: Logistic Regression

In [12] the authors investigated additive interaction with continuous exposures by modeling RERI with logistic regression. The intervention on an exposure G with some constant g_d was defined as $I_G(g_0) = g_0 + g_d$. The logistic model was

$$\frac{p_{g,e}}{1 - p_{g,e}} = e^{\alpha + \beta_1 g + \beta_2 e + \beta_3 ge}, \tag{3.43}$$

and their RERI, based on the intervention, was derived as

$$\begin{aligned} \text{RERI}^{\text{LR}}_{g,e} = {} & e^{\beta_1 g_d + \beta_2 e_d + \beta_3((g_0+g_d)(e_0+e_d) - g_0 e_0)} \\ & - e^{g_d(\beta_1 + \beta_3 e_0)} - e^{e_d(\beta_2 + \beta_3 g_0)} + 1. \end{aligned} \tag{3.44}$$

Let us now examine this model for continuous additive interaction using Theorem 3.3. Using the rare disease assumption then for the requirements of Theorem 3.3 to be met Equation 3.43 has to be able to be written in the form of Equation 3.37. This is not possible since exponential functions are multiplicative. It then follows that the model can not show no exposure additive interaction if there are marginal effects, i.e., $\beta_1 \neq 0$ and $\beta_2 \neq 0$.

The model can also not fulfill criterion (ii) since with $\beta_1 = 0$, $\beta_2 = 0$ the odds ratio for the marginal effect depend on β_3, for instance the marginal effect for G is:

$$\text{OR}^{g_0,e_0}_{g_0+g_d,e_0} = e^{\beta_3 g_d e_0} \tag{3.45}$$

Model Example: Linear Odds

A model that does fulfill the requirements for criterion (i) is the linear odds model,

$$\frac{p_{g,e}}{1-p_{g,e}} = a + b_1 g + b_2 e + b_3 ge. \tag{3.46}$$

In contrast to the logistic model the linear odds model can be written on the form of Equation 3.37, meaning that it can show no exposure additive interaction with no interaction corresponding to $b_3 = 0$. The linear odds model leads to $\text{RERI}^{\text{LO}}_{g,e}$ and $\text{AP}^{*\text{LO}}_{g,e}$ as:

$$\text{RERI}^{\text{LO}}_{g,e} = \frac{b_3(I_G(g_0) - g_0)(I_E(e_0) - e_0)}{a + b_1 g_0 + b_2 e_0 + b_3 g_0 e_0} \tag{3.47}$$

$$\text{AP}^{*\text{LO}}_{g,e} = \frac{b_3(I_G(g_0) - g_0)(I_E(e_0) - e_0)}{b_1(I_G(g_0) - g_0) + b_2(I_E(e_0) - e_0) + b_3(I_G(g_0)I_E(e_0) - g_0 e_0)} \tag{3.48}$$

The presence or direction of interaction in this model can be tested by testing if b_3 deviates from 0. However, as shown in Theorem 3.4, models of the form in Equation 3.37 will have issues with criterion (ii), that is no marginal effects and the presence of additive interaction at the same time. In the linear odds model we can see this by setting the coefficients corresponding to the marginal effects to zero, i.e., $b_1 = 0$, $b_2 = 0$, for $\text{AP}^{*\text{LO}}_{g,e}$ which becomes

$$\text{AP}^{*\text{LO}}_{g,e} = \frac{(I_G(g_0) - g_0)(I_E(e_0) - e_0)}{(I_G(g_0)I_E(e_0) - g_0 e_0)}. \tag{3.49}$$

For there to be only interaction effects and no marginal effects then it should be true that $\text{AP}^{*\text{LO}}_{g,e} = 1$ for all values of g and e, which is clearly not true unless $I_G(g_0) = I_E(e_0) = 0$ or $g_0 = e_0 = 0$. It also holds if the intervention functions are $I_G(g) = I_E(e) = 0$ for all g and e, which would mean that the interventions removes the exposures, and the linear odds model with these interventions does fulfill both criteria.

4 Discussion

The purpose of this paper is to explore models for additive interaction with continuous exposures. We have shown that with the intervention defined as a monotonic function there is no model that can fulfill both of the following criteria:

i. There exist some values for the model parameters for which both exposures have marginal effects and no additive interaction on the exposure level for both exposures.

ii. There exist some values for the model parameters for which there are no marginal effects and there is additive interaction between the exposures.

Our finding means that the previously suggested methods in the literature ([12, 13, 31]) for additive interaction and continuous exposures do not work in practice, since those methods are based on logistic regression and are unable to fulfill any of the two criteria. If there are marginal effects, those model's estimates will always conclude there is additive interaction, given that there is sufficient statistical power.

Note that the results are dependent on the definition of the intervention on the continuous exposures. The definition of the intervention is more complicated than in the binary case. Partially because the groups that can be intervened upon can be more complex, and partially because the interventions themselves can be modeled in more detail, and what an intervention is can also have different implications. In this paper we defined the intervention as a monotonic function on the exposure, other definitions could lead to different conclusions. In the linear odds example in Section 3.2.3 with the intervention defined as setting the exposure to zero both criteria do hold.

However, our results do not change the fact that the risk differences are an important part of causal inference and that the IC is the measure of interest when comparing interventions on different groups. With the correct context and model for the estimation of the odds or probabilities, the IC is still be useful.

Mechanistic interpretations of interaction with continuous exposures is different though, and would require a different causality model than the sufficient cause model, since the sufficient cause model is binary and there is no obvious way to extend the model that is not simply a transformation from continuous to binary. A continuous deterministic causality model will always have some corresponding binary sufficient cause model, since a deterministic continuous causality model can always be transformed into a binary sufficient cause model given the cutoffs for the transformation from continuous to binary is set properly, such as done in [3, 35]. The most trivial model transformation is to use the same function as for the continuous model itself, i.e., the new binary variable is set so that it is one for all values of the continuous exposures where the outcome is one in the continuous causality model and zero for all other areas. However, such a model is not trivial to estimate in practice since the true cutoffs are unknown [3, 35] Another motivation is that a deterministic model could not fulfill some of the properties that would be useful to take advantage of the extra information provided by the continuous exposures, e.g., a model where the probability increases proportionally with the exposure. However, this continuous non-deterministic causality model is not the same kind of stochastic model used in [3, 32, 35] since the exposures themselves do not have a random effect on the outcome in those models.

An investigator in a study about additive interaction with continuous exposures would need to be careful about the choice of model for the probabilities or odds, since both the aforementioned criteria can not be true for the same model. For instance if the interest of the study is masking effects then criterion (ii) is important since it tells us that one effect is masking the other.

Even though the linear odds model does fulfill criterion (i), it can be unsuitable in practice because convergence problems as explained in Section 3.2, or because a linear

increase of the risk might be unrealistic. However, in some cases a linear risk model can be realistic. One example is the linear no-threshold model for cancer induced by radiation which models the cancer risk as proportional to the amount of radiation [26].

With more general interventions, and also possibly comparisons between different possible intervention functions, since no model can give a simple answer as in the binary case using decision theory could be useful for finding the best intervention and which group to target, although to a large extent such methods would likely depend on the study's context.

References

[1] T. Andersson, L. Alfredsson, H. Källberg, S. Zdravkovic, and A. Ahlbom. Calculating measures of biological interaction. *European Journal of Epidemiology*, 20(7):575–579, 2005.

[2] S. F. Assmann, D. W. Hosmer, S. Lemeshow, and K. A. Mundt. Confidence intervals for measures of interaction. *Epidemiology*, pages 286–290, 1996.

[3] C. Berzuini and A. P. Dawid. Stochastic mechanistic interaction. *Biometrika*, 103(1):89–102, 2016.

[4] W. J. Blot and N. E. Day. Synergism and interaction: are they equivalent? *American Journal of Epidemiology*, 110(1):99–100, 1979.

[5] S. Greenland. Additive risk versus additive relative risk models. *Epidemiology*, pages 32–36, 1993.

[6] S. Greenland. Basic problems in interaction assessment. *Environmental Health Perspectives*, 101(Suppl 4):59, 1993.

[7] S. Greenland, J. M. Robins, and J. Pearl. Confounding and collapsibility in causal inference. *Statistical Science*, pages 29–46, 1999.

[8] P. W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, 1986.

[9] D. W. Hosmer and S. Lemeshow. Confidence interval estimation of interaction. *Epidemiology*, 3(5):452–456, 1992.

[10] O. Hössjer, I. Kockum, L. Alfredsson, A. K. Hedström, T. Olsson, and M. Lekman. A general framework for and new normalization of attributable proportion. *Epidemiologic Methods*, 6(1), 2017.

[11] H. Källberg, B. Ding, L. Padyukov, C. Bengtsson, J. Rönnelid, L. Klareskog, L. Alfredsson, E. S. Group, et al. Smoking is a major preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. *Annals of the Rheumatic Diseases*, 70(3):508–511, 2011.

[12] M. Katsoulis and C. Bamia. Additive interaction between continuous risk factors using logistic regression. *Epidemiology*, 25(3):462–464, 2014.

[13] M. J. Knol, I. van der Tweel, D. E. Grobbee, M. E. Numans, and M. I. Geerlings. Estimating interaction on an additive scale between continuous determinants in a logistic regression model. *International Journal of Epidemiology*, 36(5):1111–1118, 2007.

[14] M. J. Knol, J. P. Vandenbroucke, P. Scott, and M. Egger. What do case-control studies estimate? Survey of methods and assumptions in published case-control research. *American Journal of Epidemiology*, 168(9):1073–1081, 2008.

[15] M. J. Knol and T. J. VanderWeele. Recommendations for presenting analyses of effect modification and interaction. *International Journal of Epidemiology*, 41(2):514–520, 2012.

[16] M. J. Knol, T. J. VanderWeele, R. H. Groenwold, O. H. Klungel, M. M. Rovers, and D. E. Grobbee. Estimating measures of interaction on an additive scale for preventive exposures. *European Journal of Epidemiology*, 26(6):433–438, 2011.

[17] M. L. Petersen, K. E. Porter, S. Gruber, Y. Wang, and M. J. van der Laan. Diagnosing and responding to violations in the positivity assumption. *Statistical Methods in Medical Research*, 21(1):31–54, 2012.

[18] D. H. Rehkopf, M. M. Glymour, and T. L. Osypuk. The consistency assumption for causal inference in social epidemiology: When a rose is not a rose. *Current Epidemiology Reports*, 3(1):63–71, 2016.

[19] D. B. Richardson and J. S. Kaufman. Estimation of the relative excess risk due to interaction and associated confidence bounds. *American Journal of Epidemiology*, 169(6):756–760, 2009.

[20] K. J. Rothman, S. Greenland, and T. L. Lash. *Modern epidemiology*. Lippincott Williams & Wilkins, 2008.

[21] K. J. Rothman, S. Greenland, and A. M. Walker. Concepts of interaction. *American Journal of Epidemiology*, 112(4):467–470, 1980.

[22] P. Royston, D. G. Altman, and W. Sauerbrei. Dichotomizing continuous predictors in multiple regression: a bad idea. *Statistics in Medicine*, 25(1):127–141, 2006.

[23] R. Saracci. Interaction and synergism. *American Journal of Epidemiology*, 112(4):465–466, 1980.

[24] A. Skrondal. Interaction as departure from additivity in case-control studies: a cautionary note. *American Journal of Epidemiology*, 158(3):251–258, 2003.

[25] J. Splawa-Neyman, D. M. Dabrowska, T. P. Speed, et al. On the application of probability theory to agricultural experiments. *Statistical Science*, 5(4):465–472, 1990. Reprint.

[26] M. Tubiana, L. E. Feinendegen, C. Yang, and J. M. Kaminski. The linear no-threshold relationship is inconsistent with radiation biologic and experimental data. *Radiology*, 251(1):13–22, 2009.

[27] T. J. VanderWeele. Reconsidering the denominator of the attributable proportion for interaction. *European Journal of Epidemiology*, 28(10):779–784, 2013.

[28] T. J. VanderWeele. *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press, 2016.

[29] T. J. VanderWeele, Y. Chen, and H. Ahsan. Inference for causal interactions for continuous exposures under dichotomization. *Biometrics*, 67(4):1414–1421, 2011.

[30] T. J. VanderWeele and M. J. Knol. Remarks on antagonism. *American Journal of Epidemiology*, 173(10):1140–1147, 2011.

[31] T. J. VanderWeele and M. J. Knol. A tutorial on interaction. *Epidemiologic Methods*, 3(1):33–72, 2014.

[32] T. J. VanderWeele and T. S. Richardson. General theory for interactions in sufficient cause models with dichotomous exposures. *Annals of Statistics*, 40(4):2128, 2012.

[33] T. J. VanderWeele and J. M. Robins. The identification of synergism in the sufficient-component-cause framework. *Epidemiology*, 18(3):329–339, 2007.

[34] T. J. VanderWeele and J. M. Robins. Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika*, 95(1):49–61, 2008.

[35] T. J. VanderWeele and J. M. Robins. Stochastic counterfactuals and stochastic sufficient causes. *Statistica Sinica*, 22(1):379, 2012.

[36] T. J. VanderWeele and S. Vansteelandt. A weighting approach to causal effects and additive interaction in case-control studies: marginal structural linear odds models. *American Journal of Epidemiology*, 174(10):1197–1203, 2011.

[37] S. Wachholder. Binomial regression in GLIM: Estimating risk ratios and risk differences. *American Journal of Epidemiology*, 123(1):174–184, 1986.

[38] L. N. Yelland, A. B. Salter, and P. Ryan. Relative risk estimation in randomized controlled trials: a comparison of methods for independent observations. *The International Journal of Biostatistics*, 7(1):1–31, 2011.

[39] G. Y. Zou. On the estimation of additive interaction by use of the four-by-two table and beyond. *American Journal of Epidemiology*, 168(2):212–224, 2008.