Published in final edited form as:

Ann Hum Genet. 2011 January ; 75(1): 146–156. doi:10.1111/j.1469-1809.2010.00600.x.

Tests for Compositional Epistasis under Single Interaction-Parameter Models

Tyler J. VanderWeele and Nan M. Laird

Harvard School of Public Health

Abstract

Compositional epistasis is said to be present when the effect of a genetic factor at one locus is masked by a variant at another locus. Although such compositional epistasis is not equivalent to the presence of an interaction in a statistical model, non-standard tests can sometimes be used to detect compositional epistasis. In this paper we consider empirical tests for compositional epistasis under models for the joint effect of two genetic factors which place no restrictions on the main effects of each factor but constrain the interactive effects of the two factors so as to be captured by a single parameter in the model. We describe the implications of these tests for cohort, case-control, case-only and family-based study designs and we illustrate the methods using an example of gene-gene interaction already reported in the literature.

Introduction

Several authors (Cordell, 2002, 2009; Moore & Williams, 2005, 2009; Phillips, 2008) have recently distinguished “statistical epistasis” from more biologic forms of epistasis in the sense of masking or in the physical interaction of proteins. Statistical epistasis is generally simply conceived of as a departure from additivity between the effects of two genetic factors in a statistical model so that correct specification of the model requires gene-gene interaction terms in the model. Cordell (2002) noted that such statistical epistasis was quite distinct from epistasis in the sense of the masking of the effect of one genetic factor by another, as Bateson (1909) had initially conceived of the term. Phillips (2008) suggested the term “compositional epistasis” to indicate epistasis in the sense of masking and the terminology has been adopted by other authors (Cordell, 2009; Moore & Williams, 2009; VanderWeele 2010a). Phillips (2008) further noted that an additional distinction could be drawn between compositional epistasis in the sense of masking and what he called “functional epistasis” conceived of as the actual physical interaction of proteins. VanderWeele (2010ab) noted that although other authors (Cordell, 2002, 2009; Cordell & Clayton, 2005) had pointed out that standard tests for statistical interaction or statistical epistasis could not be used to draw conclusions about compositional epistasis, one could in fact use alternative non-standard tests to, in some cases, empirically test for certain forms of compositional epistasis. Although the tests derived in VanderWeele (2010ab) allowed for stronger conclusions about compositional epistasis (rather than merely statistical epistasis), the tests generally required larger effect sizes and sample sizes and consequently power becomes a considerable concern with these tests.

Power is a concern with interaction tests more generally; the literature on power and sample size calculations for interaction tests indicates that considerably larger sample sizes are often needed to detect interactions than to detect main effects (Gauderman, 2002; Wang and Zhao, 2003) and these concerns about power are amplified in the context of multiple comparisons and GWAS gene-gene interaction testing (Kraft, 2004; Musani et al., 2007; Kooperberg & LeBlanc, 2008; Pierce & Ahsan, 2010). Much of the current literature on power in the context of interactions concerns trying to leverage the presence of interactions to detect main

effects (Chatterjee et al., 2006; Kraft et al., 2007; Maity et al., 2009). However, power concerns become a yet greater issue when we specifically want to estimate the interaction parameters themselves, especially when the statistical models used to test for statistical interaction as departure from additivity in the effects of the two factors allow for complete flexibility in model parameterization. With two genetic factors coded as variables with three levels indicating 0, 1, or 2 variant alleles, a saturated model would involve five parameters for the baseline genetic risk and the main effects and four additional parameters for the interaction (Cordell, 2002). With four separate interaction parameters, power to detect statistical interaction becomes even more problematic. In order to partially circumvent this issue, some authors (Hoffmann et al., 2009; Barhdadi and Dubé, 2010) have proposed the use of models that allow for fully flexible main effects but constrain the interactive effects by e.g. requiring that it be captured by a single parameter; this allows for more efficient tests of the presence of a statistical interaction while avoiding the possibility that misspecification of the model for the main effects results in erroneously concluding a departure from additivity of the main without an interaction in fact being present.

In this paper we will use of the results of VanderWeele (2010ab) to explore the conclusions concerning compositional epistasis that can be drawn from such single interaction-parameter models when these models are in fact correctly specified. Particularly simple results concerning compositional epistasis arise under the use of such models. We will give a counterfactual exposition of compositional epistasis as in VanderWeele (2010b); we will then describe the class of single interaction-parameter models that we consider in this paper and we will discuss the implications of the use of such models for tests for compositional epistasis in cohort, case-control, case-only and family-based study designs. We conclude with an illustration and some further discussion.

Counterfactual Conception of Compositional Epistasis

Following the exposition in Cordell (2002) of what has since come to be called “compositional epistasis,” VanderWeele (2010b) related this concept of compositional epistasis to the counterfactual or potential outcomes framework (Rubin, 1990; Hernán, 2004) that has become widespread within statistics and epidemiology. Suppose that at each loci A and B there are three distinct relevant genotypes: a/a , a/A and A/A at locus A and b/b , b/B and B/B at locus B. Let G_1 and G_2 be variables with three levels indicating the genotype at loci A and B respectively (e.g. $G_1=0$ for a/a , $G_1=1$ for a/A , $G_1=2$ for A/A and $G_2=0$ for b/b , $G_2=1$ for b/B , $G_2=2$ for B/B). Let D be a binary indicator of phenotype, indicating the presence of some dichotomous trait. For each individual in the population let D_{ij} denote what the trait would have been if G_1 were i and if G_2 were j . For each individual we could conceive of what might have happened to that individual had the genotype at each locus been something other than it was. In particular we might consider whether there were any individuals in the population that had response patterns like any of those in Table 1.

Each of the response patterns in Tables 1a-1d would constitute an instance of “compositional epistasis” because, for example, the effect of the genetic factor at locus A is masked when locus B is of the b/b genotype. Note that for complex traits with non-Mendelian inheritance, the response patterns may vary from one individual to another. In some cases it may be known a priori that an increase in the number of variant alleles will never for any individual prevent the outcome so that for every individual D_{ij} is non-decreasing in i or j . We will say that G_1 has a monotonic effect on D if D_{ij} is non-decreasing in i and that G_2 has a monotonic effect on D if D_{ij} is non-decreasing in j . As will be seen below, monotonicity assumptions of this sort will more easily allow for the detection of compositional epistasis. We note, however, that monotonicity assumptions are strong assumptions insofar as they make reference to all individuals in the population. Empirical

data can sometimes be used to invalidate such monotonicity assumptions but such assumptions can never be empirically verified with data since the monotonicity assumptions make reference to all of the potential outcomes for each particular individual in a population under each possible combination of the factors and we only observe the outcome D under one particular setting of G_1 and G_2 . One would thus generally have to rely on knowledge of the biology itself to reasonably make these monotonicity assumptions. In some settings this may be possible; for example, it is difficult to imagine that mutations of the BRCA1 gene will ever be protective for breast cancer for any individual. However, in many settings, our knowledge of how precisely genetic variants might influence biological systems is likely insufficient to be able to make monotonicity assumptions with confidence.

VanderWeele (2010b) considered empirical tests for compositional epistasis of the forms in Table 1, both with and without monotonicity assumptions, using probabilities of the form $p_{ij} = P(D=1|G_1=i, G_2=j)$ i.e. using the probabilities of the outcome amongst individuals with $G_1=i, G_2=j$. For example, it was shown that if both G_1 and G_2 have monotonic effects on the outcome and if $p_{22} - p_{21} - p_{12} + p_{11} > 0$ then there must be some individuals in the population with the response pattern like that in Table 1a (i.e. instances of compositional epistasis). If only G_1 say has a monotonic effect on the outcome then to detect instances of compositional epistasis of the form in Table 1a one could test $p_{22} - p_{21} - p_{20} - p_{12} > 0$. Tests for other forms of compositional epistasis in Table 1 and for settings when no assumptions are made about monotonicity were also given in VanderWeele (2010b).

The contribution in this paper over prior work on empirical tests for compositional epistasis is three-fold: first, whereas prior work on empirical tests for compositional epistasis (VanderWeele, 2010ab) only considered tests for complete response patterns (as in Table 1), it is noted here that instances of compositional epistasis, under the counterfactual conception, can be detected even when part of the response pattern is unknown (see Table 2 below) and we provide tests for such instances of compositional epistasis. Second, we provide an extensive characterization of tests for compositional epistasis in linear and log-linear/logistic single interaction parameter models; this characterization will facilitate the application of tests for compositional epistasis in practice and will be useful in increasing power for such tests when single interaction parameter models fit the data. Third, we consider inference about compositional epistasis in family-based study designs.

Statistical Models with a Single Interaction Parameter

As above, we let $p_{ij} = P(D=1|G_1=i, G_2=j)$ denote the probability of the outcome amongst individuals with $G_1=i, G_2=j$ i.e. the penetrance for $G_1=i, G_2=j$. We use the notation $1(V=v)$ to be a function that takes the value 1 if $V=v$ and 0 otherwise. Under the setting considered above, a fully general model for penetrance probabilities would be

$$P(D=1|G_1=g_1, G_2=g_2) = \mu + \alpha_1 1(g_1=1) + \alpha_2 1(g_1=2) + \beta_1 1(g_2=1) + \beta_2 1(g_2=2) + \lambda_{11} 1(g_1=1)1(g_2=1) + \lambda_{21} 1(g_1=2)1(g_2=1) + \lambda_{12} 1(g_1=1)1(g_2=2) + \lambda_{22} 1(g_1=2)1(g_2=2) \quad (1)$$

The model given in equation (1) does not impose any restrictions on the data but contains four separate interaction parameters λ_{11} , λ_{21} , λ_{12} , λ_{22} , and large sample sizes may be required to be able to detect any statistically significant interaction at all. To attempt to partially address these issues of power to detect interactions, we will instead consider statistical models that impose no assumptions on the main effects but involve only a single interaction parameter and take the form

$$P(D=1|G_1=g_1, G_2=g_2)=\mu+\alpha_1 1(g_1=1)+\alpha_2 1(g_1=2)+\beta_1 1(g_2=1)+\beta_2 1(g_2=2)+\lambda_{\text{int}} g_1 g_2 \quad (2)$$

Note that the final term $g_1 g_2$ takes the value 0 if $G_1=0$ or $G_2=0$, takes the value 1 if $G_1=G_2=1$, takes the value 2 if one of G_1 or G_2 is 1 and the other is 2, and takes the value 4 if $G_1=G_2=2$. Model (2) imposes structure and restrictions on the form of the interaction but as a result allows for interaction to be captured using a single interaction parameter λ_{int} , rather than four interaction parameters, λ_{11} , λ_{21} , λ_{12} , λ_{22} . The parameters of model (2) could be fit by maximum likelihood using standard statistical software for fitting generalized linear models. Model (2) falls within the class of AMMI models considered by Barhdadi and Dubé (2010). See also Song & Nicolae (2010) for other types of gene-gene interaction models with restrictions on the parameter space. As discussed below in the section on study design, we note that linear models with an identity link, such as (1) and (2) can be fit, up to a constant of proportionality μ , even with case-control data.

Model (2) also allows for fairly simple tests for compositional epistasis as in Tables 1a-1d. For now we will assume that the penetrance probabilities $p_{ij} = P(D=1|G_1=i, G_2=j)$ reflect the true effects of genetic factors G_1 and G_2 ; in a subsequent section we will consider how tests for compositional epistasis can be adapted to control for possible population stratification or confounding. Derivations of the following results are given in the online supplementary materials. Suppose first that both G_1 and G_2 have monotonic effects on D then if $\lambda_{\text{int}} > 0$ this implies the presence of at least some individuals with response patterns such as that of Table 1a i.e. compositional epistasis is present for at least some individuals. If it is in fact the case that $\lambda_{\text{int}} > (\alpha_2 - \alpha_1) + (\beta_2 - \beta_1)$ then there are individuals with response pattern of Table 1d. Under model (2), one can also sometimes detect forms of compositional epistasis even if $\lambda_{\text{int}} < 0$. It can be shown that if $\lambda_{\text{int}} > (\beta_1 - \alpha_1)$ then there are individuals with response pattern given in Table 2a (where the '?' in Table 2 denotes values that could be 0 or 1); if $\lambda_{\text{int}} > (\alpha_1 - \beta_1)$ then there are individuals with response pattern given in Table 2b; provided α_1 and β_1 are not equal, one of $(\beta_1 - \alpha_1)$ or $(\alpha_1 - \beta_1)$ will be negative.

Note that the response pattern in Table 2a implies compositional epistasis since the effect of the genetic factor at locus B (evident when the genotype at locus A is a/A) is masked when the genotype at locus A is a/a. Similar remarks hold for Table 2b where the effect of the genetic factor at locus A (evident when the genotype at locus B is b/B) is masked when the genotype at locus B is b/b. Note also that Table 2a is consistent with the epistatic response patterns given in Tables 1b and 1d (and others) and Table 2b with the response patterns given in Tables 1c and 1d (and others).

The tests we have just described made the assumption that the effects of both G_1 and G_2 on D are monotonic. This is a strong assumption and in many contexts will not hold. We can also consider tests for compositional epistasis when only one, or when neither of G_1 and G_2 , have monotonic effects on the outcome D . These will require more stringent statistical tests; without monotonicity of both factors, a positive value of λ_{int} will, on its own, no longer suffice to conclude the presence of compositional epistasis. These further tests, along with the tests described above, are summarized in Table 3 which lists (i) the assumptions about monotonicity needed for the test, (ii) the condition to be tested expressed in terms of the coefficients of model (2) and (iii) the form of compositional epistasis which must be present if the condition is satisfied.

For example, suppose that only one of the genetic factors, say G_1 , has a monotonic effect on D then, as reported in the fifth line of the table, if $\lambda_{\text{int}} > \mu/4$ then there are at least some individuals with response pattern given in Table 2c which once again implies compositional epistasis since the effect of the genetic factor at locus A (evident when the genotype at locus

B is B/B) is masked when the genotype at locus B is b/b. Note that if model (2) is indeed correctly specified then when only one or neither of the genetic factors have a monotonic effects on the outcome then it is not possible to test for compositional epistasis of the forms in Table 1. This does not mean that such epistatic response patterns are not present, only that it will not be possible to detect them by statistical tests. The conditions in Table 3 and all subsequent tables are sufficient conditions for compositional epistasis but not necessary conditions.

The conditions in Table 3 can also be used to estimate lower bounds on the prevalence of individuals manifesting response patterns which constitute instances of compositional epistasis. Specifically, the difference between the left side and the right side of the inequalities in the second column of Table 3 give lower bounds on the prevalence of the corresponding form of compositional epistasis. Thus for example, on the fifth line of Table 3 (assuming the effect of G_1 is monotonic), $\lambda_{\text{int}} - \mu/4$ gives a lower bound on the proportion of individuals that manifest epistasis of the form indicated in Table 2c. Similar remarks apply to all results concerning linear models with identity links in this paper. For further discussion on prevalence bounds, see VanderWeele et al. (2010a).

We have seen then that $\lambda_{\text{int}} > 0$ in model (2) only necessarily implies compositional epistasis under the strong assumption that both G_1 and G_2 have monotonic effects on D. However, even when this assumption is violated we can still test for compositional epistasis but we need stronger statistical tests i.e. more stringent conditions for λ_{int} need to be satisfied.

Statistical Models with a Logit Link and a Single Interaction Parameter

In many analyses with dichotomous outcomes and in many case-control studies, rather than using a model with a linear link like (1) or (2) above, logistic regression models (i.e. models with a logit link) are used instead. In this section we will consider tests for compositional epistasis such as is present in Tables 1 and 2 above in the context of single interaction-parameter statistical models with logit links. The analogous model to (2) with a logit link is:

$$\text{logit}(p(D=1|G_1=g_1, G_2=g_2)) = \mu^\dagger + \alpha_1^\dagger I(g_1=1) + \alpha_2^\dagger I(g_1=2) + \beta_1^\dagger I(g_1=1) + \beta_2^\dagger I(g_1=2) + \gamma_{\text{int}}^\dagger g_1 g_2$$

where we use $\mu^\dagger, \alpha_1^\dagger, \alpha_2^\dagger, \beta_1^\dagger, \beta_2^\dagger$, rather than $\mu, \alpha_1, \alpha_2, \beta_1, \beta_2$, so as to be able to distinguish between the parameters in model (2) with the identity link from those of model (3) with the logit link. The parameters of model (3) could be fit by maximum likelihood using standard statistical software for fitting logistic regression models. We will assume throughout that the outcome D is relatively rare under all combinations of G_1 and G_2 so that odds ratios approximate risk ratios and the logit link approximates a log link.

Tests for compositional epistasis somewhat analogous to those for model (2) can also be used for model (3). Assume that penetrance probabilities $p_{ij} = P(D=1|G_1=i, G_2=j)$ are non-decreasing in i and j even if the assumption of individual level monotonic effects (that D_{ij} is non-decreasing in i and j for every individual) does not hold, then the conditions listed in the second column of Table 4, along with the monotonicity assumption in the first column allow one to conclude the presence of the form of compositional epistasis listed in the third column.

As was the case in the model with identity link, the fewer the assumptions made about monotonicity, the stronger conditions are needed on γ_{int} in order to conclude the presence of compositional epistasis.

Settings in which one of the factors is dichotomous

Suppose now that G_1 has three levels but that G_2 can effectively be considered binary either because B/B genotype has frequency of 0 or because the mode of inheritance for G_2 is known a priori to be recessive (in which case $G_2=0$ for the b/b and b/B genotype and $G_2=1$ for B/B) or dominant (in which case $G_2=0$ for the b/b genotype and $G_2=1$ for b/B or B/B). Results when both factors are dichotomous are given in VanderWeele (2010a). We again let D_{ij} denote, for each individual in the population, what the dichotomous trait D would have been if G_1 were i and if G_2 were j and we let $p_{ij} = P(D=1|G_1=i, G_2=j)$ denote the penetrance for $G_1=i, G_2=j$. Various forms of compositional epistasis in this setting are presented in Table 5. Note that all of the response patterns in Tables 5a-5d manifest compositional epistasis because the effect of the genetic factor at locus A is masked when $G_2=0$.

A statistical model with linear link which places no restrictions on the main effects but has a single interaction parameter is given by:

$$P(D=1|G_1=g_1, G_2=g_2) = \mu + \alpha_1 1(g_1=1) + \alpha_2 1(g_1=2) + \beta_1 1(g_2=1) + \lambda_{\text{int}} g_1 g_2 \quad (4)$$

where the final term $g_1 g_2$ takes the value 0 if $G_1=0$ or $G_2=0$, takes the value 1 if $G_1=G_2=1$ and takes the value 2 if $G_1=2, G_2=1$. Tests for various forms of compositional epistasis, expressed in terms of the coefficients of model (4), under various monotonicity assumptions are presented in Table 6.

A statistical model with logistic link which places no restrictions on the main effects but has a single interaction parameter is given by:

$$\text{logit}(P(D=1|G_1=g_1, G_2=g_2)) = \mu^\dagger + \alpha_1^\dagger 1(g_1=1) + \alpha_2^\dagger 1(g_1=2) + \beta_1^\dagger 1(g_2=1) + \gamma_{\text{int}} g_1 g_2 \quad (5)$$

Assume that the outcome is rare for all combinations of G_1 and G_2 and that the penetrance probabilities $p_{ij} = P(D=1|G_1=i, G_2=j)$ are non-decreasing in i and j then Table 7 gives results conditions for compositional epistasis under model (5).

Cohort, Case-Control, Case-Only and Family-Based Study Designs

In this section we will consider how the tests for compositional epistasis described above could be employed in a variety of study designs. In the remainder of the paper, we will restrict our discussion to the setting in which both factors have three levels as in models (2) and (3) and Tables 3 and 4. However, similar remarks apply also when one of the factors has only two levels.

In cohort studies we could fit models (2) or (3) and obtain estimates of all of the parameters and could thus apply any of the tests for compositional epistasis considered above. In a case-control study, model (2) cannot be fit unless data is available on the prevalence of disease (Rothman et al., 2008). However, case-control data can be used to fit model (3) to obtain estimates of all parameters in model (3) except $\mu^\dagger$. None of the tests described above using γ_{int} from model (3) required $\mu^\dagger$ and thus all of these tests could be applied when using case-control data. Although model (2) cannot be fit using case-control data, provided the outcome is rare for all combinations of G_1 and G_2 so that odds ratios approximate risk ratios and the logit link approximates a log link, all of the parameters of model (2) could be estimated up through a proportionality constant $\mu = p_{00}$; that is to say, each of α_1/μ , α_2/μ , β_1/μ , β_2/μ and λ_{int}/μ could be estimated from case-control data. Consequently, one could still test the conditions given in the section on single interaction-parameter models with identity link by

estimating each of α_1/μ , α_2/μ , β_1/μ , β_2/μ and λ_{int}/μ using case-control data and then dividing both sides of the inequalities in Table 3 by μ . A similar approach is often used in epidemiologic research to obtain measures of interaction on an additive scale using case-control data often described as the “relative excess risk due to interactions” or “RERI” (Rothman, 1986).

We show in the online supplementary materials that γ_{int} from model (3) can be estimated from case-only data (Piergorsch et al., 1994) provided that the two genetic factors are independent in the population (as would usually hold if the two genetic factors were on different chromosome) and that the outcome is rare for all combinations of G_1 and G_2 so that odds ratios approximate risk ratios and the logit link approximates a log link. However, with case-only data, none of the other parameters in model (3) can be estimated. Thus the only tests for compositional epistasis that could be used with case-only data are those which rely only on the parameter γ_{int} . These tests generalize remarks on case-only designs in VanderWeele et al. (2010b) to settings where the genetic factors have three levels rather than being binary.

In family-based study designs based on discordant sib pairs or sibships (Witte, et al., 1999), when G_1 or G_2 are both genetic factors, then all of the parameters in model (3) except $\mu^\dagger$ can be estimated and thus all of the tests for compositional epistasis described above for model (3) could be employed in these family-based designs of gene-gene interaction. With case-parent designs, where genotype data are available on cases only plus their parents (Cordell et al., 2004) all of the parameters (apart from the intercepts) in model (3) can be estimated under the rare disease assumption. With model (2), μ cannot be estimated, but dividing both numerator and denominator by μ allows us to estimate α_1/μ , α_2/μ , β_1/μ , β_2/μ and λ_{int}/μ . Results for both models (2) and (3) require that we know the distribution of G_1 and G_2 conditional on the parental haplotypes at both loci. This is the case if the two genetic loci are not linked so that G_1 and G_2 are conditionally independent given parents, or if they are sufficiently close that we can assume that no recombination has occurred between them. In general, however, this distribution is unknown, and none of the parameters can be estimated without additional assumptions.

Control for Confounding and Population Stratification

The tests we have described above require that the penetrance probabilities reflect the true effects of the genetic factors on the outcome. This may not be the case due to population stratification or confounding. If control can be made for population stratification or confounding by means of some vector of covariates C then the tests described above could still be employed. If C contains a small number of binary or categorical variables then the tests described above could be applied within each stratum of the covariates C . If C contains continuous covariates or many categorical covariates then it may be desirable to control for confounding by incorporating the variables C into the regression model. All of the tests described above for logistic models (3) and (5) will still be applicable if a term $\delta'C$ is included in regression model, provided that the model is correctly specified; essentially regression estimates δ for C can effectively be ignored. This is because the $\delta'C$ term drops out of the probability expressions in the tests for compositional epistasis; this would not be the case if there were interactions between C and G_1 or G_2 (VanderWeele, 2009, 2010b). Of the tests described above for model (2) or (4) with identity link, only the tests given above under the assumption that both factors have monotonic effects on the outcome will be valid if a term $\delta'C$ is included in model. The tests described for compositional epistasis when only one or when neither factor has a monotonic effect could not be directly employed; this is because if a term $\delta'C$ is included in model (2) or (4) then tests for compositional epistasis will in fact then depend on the value of C . An alternative approach that can be used to

control for confounding and which can be directly applied to models (2)-(5) is one in which control for confounding is done not by regression but by using an inverse-probability-weighting technique (Robins and Hernán, 2006). VanderWeele et al. (2010a) discusses some of the relative advantages and disadvantages of regression versus weighting for confounding control in tests for interactions.

Note that both types of family designs (sibships and case-parent) control for confounding due to population substructure by using analyses conditional on parental genotype (case-parent) or conditional on family membership (sibships). Thus implicitly, the intercept term for models (2) and (3) can be replaced by $\mu(P)$ or $\mu^{\dagger}(P)$, where P indicates parental genotype, since this term drops out of the conditional likelihoods used for both designs. For the case-parent design, we can further control for individual level confounding, since $\mu(P)$ or $\mu^{\dagger}(P)$ can be replaced by $\mu(P,C)$ or $\mu^{\dagger}(P,C)$. See details in the Online Supplementary Material.

Illustration

To illustrate the methods, we will apply the tests described above to data reported by Källberg et al. (2007) who consider possible gene-gene interaction between HLA-DRB1 and R620W PTPN22 alleles on anti-CCP-Positive rheumatoid arthritis. In Table 5 of their paper, Källberg et al. (2007), report, from pooling three case-control studies, numbers of cases and controls by the presence of zero, one or two HLA-DRB1 SE alleles ($G_1=0,1,2$) and by the presence versus absence of minor R620W PTPN22 allele ($G_2=0,1$). Our analysis here is given for illustrative purposes only as it uses only the number of cases and controls reported by Källberg et al. (2007) and is not able to account for possible confounding; a full examination of evidence for possible compositional epistasis would require re-analysis of the data to control for confounding. Under a rare disease assumption, we are able to estimate the parameters of model (4) up through a proportionality constant μ ; that is, we can estimate α_1/μ , α_2/μ , β_1/μ and λ_{int}/μ (Greenland, 1993). In this case, the single interaction-parameter model (4) fits the data reasonably well. The likelihood ratio test comparing model (4) with a saturated model does not reject the null that model (4) fits the data; the AIC and BIC are also lower for model (4) than for the saturated model. Estimates for these parameters are:

$$\alpha_1/\mu = 3.75 \text{ (95\% CI: 2.73, 4.77)}$$

$$\alpha_2/\mu = 13.97 \text{ (95\% CI: 10.05, 17.88)}$$

$$\beta_1/\mu = 0.55 \text{ (95\% CI: -0.01, 1.10)}$$

$$\lambda_{\text{int}}/\mu = 5.63 \text{ (95\% CI: 3.33, 7.92)}$$

Using the results in Table 6, we can test for different forms of compositional epistasis under different assumptions. These tests are carried out in Table 8 which rearranges the condition in Table 6 so that both sides of the inequality are divided by μ . The final two columns indicate whether the conditions required to conclude the presence of each particular form of compositional epistasis is satisfied for the point estimate of the contrast in the second column and whether it is satisfied for the entire 95% confidence for the contrast.

There is evidence for compositional epistasis of the form in Table 5a when it can be assumed that the effects of both G_1 and G_2 are monotonic (i.e. both HLA-DRB1 SE and minor R620W PTPN22 alleles have monotonic effects on the outcome) since the entire confidence interval (3.3,7.9) satisfies the condition needed to conclude compositional epistasis of the form in Table 5a. However, under weaker assumptions about monotonicity, we do not have much evidence to conclude this form of compositional epistasis. Under the assumption that just G_1 is monotonic (i.e. just that HLA-DRB1 SE alleles have monotonic effects), although the point estimate of the contrast $\lambda_{\text{int}}/\mu - (\alpha_1/\mu - 1) = 0.9$ would still give

evidence for compositional epistasis of the form in Table 5a, the confidence interval for this contrast, (-1.4, 3.2), includes 0. Interestingly, however, there is evidence for compositional epistasis of the form of Table 5c irrespective of monotonicity assumption. Even without any assumptions on the monotonicity of the two genetic factors, the estimate and the confidence interval (2.3; 95% CI: 0.0, 4.5) suggest that this form of compositional is present. This is of particular interest in that, until more is understood about the biological role of the genetic variants considered, it is probably best not to make monotonicity assumptions. In this example, the gain in power by using a single interaction parameter model is important. If, instead of employing such a model along with the tests described in this paper, we use the empirical tests for compositional described in VanderWeele (2010b), without making modeling assumptions, then of the eleven conditions considered in Table 8, the only test which would provide statistically significant evidence for compositional epistasis is that for the form of compositional epistasis of Table 5c under the assumption that at least the effect of G_1 is monotonic (i.e. only for the fifth line in Table 8 is there statistically significant evidence of compositional epistasis). In particular then, without using a single interaction parameter model, we could not draw conclusions about compositional epistasis of any form without assumptions about monotonicity.

Once again, these conclusions presuppose that the associations between HLA-DRB1 SE and minor R620W PTPN22 alleles on rheumatoid arthritis reflect actual effects and are not confounded; a more reliable assessment would involve reanalyzing the data to control for possible confounding; the results here are included for illustrative purposes only.

Discussion

The principal limitation of the tests for compositional epistasis that we have described in this paper is that they require that the single interaction-parameter model, such as models (2)-(5), is correctly specified. Although the models we have discussed impose no assumptions on the main effects of either of the two genetic factors, the models do constrain the interactive effects so as to be captured by a single parameter. This may not be a reasonable assumption. Fortunately, it is an assumption that is possible to test with data. In practice one might use a likelihood ratio test to compare an unrestricted model (such as model (1) above) with the single interaction-parameter model (such as (2)). If this test rejects the null that the penetrance probabilities are captured by the single interaction-parameter model then one should not precede with the tests for compositional epistasis that we have described in this paper. More general tests for compositional epistasis that do not impose the assumption of a single parameter for interactive effects are described elsewhere (VanderWeele, 2010b). The advantage of using single interaction-parameter models when such models do fit the data is that they will have more power to detect interactions because, for example, only one parameter need be estimated rather than four. Such models are necessarily correctly specified under the null of no interactive effects. One disadvantage of using a likelihood ratio test to compare an unrestricted model with the single interaction parameter model is that the operating characteristics in terms of type I error for this two stage approach may differ from nominal rates due to ignoring the uncertainty of the first step test. Future work could consider deriving formal statistical properties for the two stage approach. Also, we have only considered particular parameterizations of penetrance probabilities that involve a single interaction parameter; other parameterizations involving only a single interaction parameter are also possible and tests for compositional epistasis for such alternative parameterizations could also be derived.

The power advantages of using these single interaction parameter models is arguably particularly relevant in the context of testing for compositional epistasis because, as we have seen above, the conditions needed to draw conclusions about compositional epistasis are in

general more stringent than those required simply to conclude the presence of a statistical interaction. This point leads us to another limitation of the results we have presented. In many cases, it may be known that variants at particular locus are associated with disease and it may be desirable to test whether any genetic variant at a large number of other loci interact with variants at the primary locus. As the number of loci one considers increases, it will be necessary to adjust for multiple testing in order to control type I error rates. Because the tests for compositional epistasis are as stringent as they are, even under single interaction parameter models, very large sample size may be needed to detect compositional epistasis in settings in which tests for numerous combinations of loci are being considered. Because of this, the applicability of approach we have described here may be best suited to settings in which a particular candidate pair of loci is already specifically in view.

A final limitation of our results as we have presented them is the counterfactual framework itself, which, for a particular individual, traditionally presupposes a deterministic outcome under each possible exposure combination (in this setting, for each possible combination of the genetic variants). In reality, the actual biological systems giving rise to particular phenotypes may be better conceptualized as stochastic with each individual having some probability of the outcome under each possible fixed combination of the genetic variants. The counterfactual framework can be reformulated in terms of stochastic counterfactuals and stochastic response patterns (Robins and Greenland, 1989, 2000). Within this setting the results we presented here would also have to be reinterpreted. Under a stochastic counterfactual setting, if the tests we have given for “compositional epistasis” were satisfied, one could then only conclude that there were individuals such that, under particular stochastic states, the effect of a genetic factor at one locus is masked by a variant at another locus. The conclusion would thus need to be modified to refer to both individuals and stochastic states rather than simply to individuals in the population.

The tests described here would also hold for tests for gene-environment interactions and we could refer to response patterns like those in Tables 1a-1d as instances of “compositional gene-environment interaction” if one of G_1 or G_2 were an environmental, rather than a genetic, factor. However, in many settings an environmental exposure will be continuous rather than having two or three categories, and in such settings the results given here would be inapplicable. Future work will consider what conclusions can be drawn when applying interaction tests or tests for compositional epistasis to a continuous exposure that has been dichotomized. When the environmental exposure does in fact have two or three categories, our comments concerning testing for such compositional response patterns in various study designs would also apply with the exception of those that were made for family-based studies. In a number of family based study designs, when gene-environment interaction is of interest the main effect for the genetic factor and the gene-environment interaction parameter can be estimated but the main effect for the environmental factor cannot be estimated without the loss of the robustness properties of the design. If a family-based design is used and tests for compositional gene-environment interaction are of interest, one could still employ the tests described above which require only estimates of γ_{int} . Alternatively, it may be possible to derive new tests for compositional gene-environment interaction in family-based studies that make use of estimates of γ_{int} and of the main effect coefficients for just the genetic but not the environmental factor. This is a topic of current research. Future research could also consider the likely sample size requirements needed to power tests for compositional epistasis in a number of genetic study designs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by NIH grant R01 ES017876.

References

- Barhdadi A, Dubé MP. Testing for gene-gene interaction with AMMI models. *Statistical Applications in Genetics and Molecular Biology*. 2010; 9:1–27. Article 2.
- Bateson, W. *Mendel's Principles of Heredity*. Cambridge University Press; Cambridge: 1909.
- Chatterjee N, Kalaylioglu Z, Moslehi R, Peters U, Wacholder S. Powerful multilocus tests of genetic association in the presence of gene-gene and gene-environment interactions. *Am J Hum Genet*. 2006; 79:1002–1016. [PubMed: 17186459]
- Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet*. 2002; 11:2463–2468. [PubMed: 12351582]
- Cordell HJ. Detecting gene-gene interaction that underlie human diseases. *Nat Rev Genet*. 2009; 10:392–404. [PubMed: 19434077]
- Cordell HJ, Clayton DG. Genetic epidemiology 3 - genetic association studies. *Lancet*. 2005; 366:1121–1131. [PubMed: 16182901]
- Cordell HJ, Barratt BJ, Clayton DG. Case/pseudocontrol analysis in genetic association studies: a unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genetic Epidemiology*. 2004; 26:167–185. [PubMed: 15022205]
- Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. *American Journal of Epidemiology*. 2002; 155:478–484. [PubMed: 11867360]
- Greenland S. Additive risk versus additive relative risk models. *Epidemiology*. 1993; 4:32–36. [PubMed: 8420578]
- Hernán MA. A definition of causal effect for epidemiological studies. *J Epidemiol Comm Health*. 2004; 58:265–271.
- Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *Journal of Epidemiology and Community Health*. 2006; 60:578–586. [PubMed: 16790829]
- Hoffmann TJ, Lange C, Vansteelandt S, Laird NM. Gene-environment interaction tests for dichotomous traits in trios and sibships. *Genet Epidemiol*. 2009; 33:691–699. [PubMed: 19365860]
- Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM. Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genetic Epidemiology*. 2004; 26:61–69. [PubMed: 14691957]
- Källberg H, Padyukov L, Plenge RM, Ronnelid J, Gregersen PK, van der Helmvan Mil AH, Toes RE, Huizinga TW, Klareskog L, Alfredsson L. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. *American Journal of Human Genetics*. 2007; 80:867–75. [PubMed: 17436241]
- Kooperberg C, LeBlanc M. Increasing the power of identifying gene $\times$ gene interactions in genome-wide association studies. *Genetic Epidemiology*. 2008; 32:255–263. [PubMed: 18200600]
- Kraft P. Multiple comparisons in studies of gene $\times$ gene and gene $\times$ environment interaction. *American Journal of Human Genetics*. 2004; 74:582–584. [PubMed: 14973784]
- Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect disease susceptibility loci. *Human Heredity*. 2007; 63:111–119. [PubMed: 17283440]
- Laird NM, Lange C. Family-based designs in the age of large-scale gene-association studies. *Nat Rev Genet*. 2006; 7:385–394. [PubMed: 16619052]
- Maity A, Carroll RJ, Mammen E, Chatterjee N. Testing in semiparametric models with interaction, with applications to gene-environment interactions. 2009; 71:75–96. *Journal of the Royal Statistical Society, Series B*
- Moore JH, Williams SM. Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. *BioEssays*. 2005; 27:637–646. [PubMed: 15892116]

- Moore JH, Williams SM. Epistasis and its implications for personal genetics. *American Journal of Human Genetics*. 2009; 85:309–320. [PubMed: 19733727]
- Musani SK, Shriner D, Liu N, Feng R, Coffey CS, Yi N, Tiwari HK, Allison DB. *Human Heredity*. 2007; 63:67–84. [PubMed: 17283436]
- Phillips PC. Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet*. 2008; 9:855–867. [PubMed: 18852697]
- Piegorsch WW, Weinberg CR, Taylor JA. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case–control studies. *Stat Med*. 1994; 13:153–162. [PubMed: 8122051]
- Pierce BL, Ahsan H. Case-only genome-wide interaction study of disease risk, prognosis and treatment. *Genetic Epidemiology*. 2010; 34:7–15. [PubMed: 19434715]
- Robins JM, Greenland S. The probability of causation under a stochastic model for individual risk. *Biometrics*. 1989; 45:1125–38. [PubMed: 2611320]
- Robins JM, Greenland S. Comment on: “Causal inference without counterfactuals” by A.P. Dawid. *J Am Statist Assoc*. 2000; 95:477–82.
- Rothman, KJ. *Modern Epidemiology*. 1st. Little, Brown and Company; Boston, MA: 1986.
- Rothman, KJ.; Greenland, S.; Lash, TL. *Modern Epidemiology*. 3rd. Philadelphia: Lippincott Williams and Wilkins; 2008.
- Rubin DB. Formal modes of statistical inference for causal effects. *Journal of Statistical Planning and Inference*. 1990; 25:279–292.
- Song M, Nicolae DL. Restricted parameter space models for testing gene-gene interaction. *Genetic Epidemiology*. 2009; 33:386–393. [PubMed: 19058263]
- VanderWeele TJ. Sufficient cause interactions and statistical interactions. *Epidemiology*. 2009; 20:6–13. [PubMed: 19234396]
- VanderWeele TJ. Empirical tests for compositional epistasis. *Nature Reviews Genetics*. 2010a; 11:166.
- VanderWeele TJ. Epistatic interactions. *Statistical Applications in Genetics and Molecular Biology*. 2010b; 9(Article 1):1–22.
- VanderWeele TJ. Sufficient cause interactions for categorical and ordinal exposures with three levels. *Biometrika*. 2010c in press.
- VanderWeele TJ, Vansteelandt S, Robins JM. Marginal structural models for sufficient cause interactions. *American Journal of Epidemiology*. 2010a; 171:506–514. [PubMed: 20067916]
- VanderWeele TJ, Hernández-Díaz S, Hernán MA. Case-only gene-environment interaction studies: when does association imply mechanistic interaction? *Genetic Epidemiology*. 2010b; 34:327–334. [PubMed: 20039380]
- Witte JS, Gauderman WJ, Thomas DC. Asymptotic bias and efficiency in case-control studies of candidate genes and gene-environment interactions: basic family designs. *American Journal of Epidemiology*. 1999; 149:693–705. [PubMed: 10206618]
- Wang S, Zhao H. Sample size needed to detect gene-gene interactions using association designs. *American Journal of Epidemiology*. 2003; 158:899–914. [PubMed: 14585768]

Table 1

Examples of Compositional Epistasis

Table 1a				Table 1b			
b/b	B/B	b/B	B/B	b/b	b/B	B/B	B/B
a/a	0	0	0	0	0	0	0
a/A	0	0	0	0	0	0	1
A/A	0	0	1	0	0	0	1

Table 1c				Table 1d			
b/b	B/B	b/B	B/B	b/b	b/B	B/B	B/B
a/a	0	0	0	0	0	0	0
a/A	0	0	0	0	0	1	1
A/A	0	1	1	0	1	0	1

Table 2

Other Examples of compositional Epistasis

Table 2a				Table 2b			
	b/b	b/B	B/B	b/b	b/B	B/B	
a/a	0	0	0	0	0	?	
a/A	0	0	1	0	0	?	
A/A	?	?	1	0	1	1	

Table 2c				Table 2d			
	b/b	b/B	B/B	b/b	b/B	B/B	
a/a	0	?	0	0	0	0	
a/A	0	?	?	?	?	?	
A/A	0	?	1	0	?	1	

Table 3
Tests for Compositional Epistasis Under Model (2)

Monotonicity Assumption	Condition on Model (2)	Form of Epistasis
G ₁ and G ₂ Monotonic	$\lambda_{int} > 0$	Table 1a
	$\lambda_{int} > (\alpha_2 - \alpha_1) + (\beta_2 - \beta_1)$	Table 1d
	$\lambda_{int} > (\beta_1 - \alpha_1)$	Table 2a
	$\lambda_{int} > (\alpha_1 - \beta_1)$	Table 2b
G ₁ Monotonic	$\lambda_{int} > \mu/4$	Table 2c
G ₂ Monotonic	$\lambda_{int} > \mu/4$	Table 2d
No Assumption	$\lambda_{int} > (\beta_1 + 3\mu)/4$	Table 2c
	$\lambda_{int} > (\alpha_1 + 3\mu)/4$	Table 2d

Table 4
Tests for Compositional Epistasis Under Model (3)

Monotonicity Assumption	Condition on Model (3)	Form of Epistasis
G ₁ and G ₂ Monotonic	$\gamma_{int} > 0$	Table 1a
	$\gamma_{int} > (\alpha_2^\dagger - \alpha_1^\dagger) + (\beta_2^\dagger - \beta_1^\dagger)$	Table 1d
	$\gamma_{int} > (\beta_1^\dagger - \alpha_1^\dagger)$	Table 2a
	$\gamma_{int} > (\alpha_1^\dagger - \beta_1^\dagger)$	Table 2b
G ₁ Monotonic	$\gamma_{int} > \log(2)/4$	Table 2c
G ₂ Monotonic	$\gamma_{int} > \log(2)/4$	Table 2d
G ₁ or G ₂ Monotonic	$\gamma_{int} > \log(3)$	Table 1a
No Assumption	$\gamma_{int} > \log(4)/4$	Table 2c and 2d
	$\gamma_{int} > \log(8)$	Table 1a

Table 5
Compositional Epistasis When One Factor is Binary

	Table 5a		Table 5b	
	G ₂ =0	G ₂ =1	G ₂ =0	G ₂ =1
a/a	0	0	0	0
a/A	0	0	0	1
A/A	0	1	0	1
	Table 5c		Table 5d	
	G ₂ =0	G ₂ =1	G ₂ =0	G ₂ =1
a/a	0	0	0	?
a/A	0	?	0	0
A/A	0	1	0	1

Table 6
Tests for Compositional Epistasis Under Model (4)

Monotonicity Assumption	Condition on Model (4)	Form of Epistasis
G ₁ and G ₂ Monotonic	$\lambda_{int} > 0$	Table 5a
	$\lambda_{int} > (\alpha_2 - \alpha_1)$	Table 5b
G ₁ Monotonic	$\lambda_{int} > \alpha_1 + \mu$	Table 5a
	$\lambda_{int} > (\alpha_2 - \alpha_1) + \mu$	Table 5b
	$\lambda_{int} > \mu/2$	Table 5c
G ₂ Monotonic	$\lambda_{int} > \alpha_1 + \beta_1 + 2\mu$	Table 5a
	$\lambda_{int} > (\alpha_1 + 2\mu)$	Table 5d
	$\lambda_{int} > (\alpha_1 + 2\mu)/2$	Table 5c
No Assumption	$\lambda_{int} > 2\alpha_1 + \beta_1 + 4\mu$	Table 5a
	$\lambda_{int} > 2\alpha_1 + 3\mu$	Table 5d
	$\lambda_{int} > (\alpha_1 + 3\mu)/2$	Table 5c

Table 7
Tests for Compositional Epistasis Under Model (5)

Monotonicity Assumption	Condition on Model (5)	Form of Epistasis
G ₁ and G ₂ Monotonic	$\gamma_{int} > 0$	Table 5a
	$\gamma_{int} > (\alpha_2^\dagger - \alpha_1^\dagger)$	Table 5b
G ₁ Monotonic	$\gamma_{int} > \log(2)$	Table 5a
	$\gamma_{int} > (\alpha_2^\dagger - \alpha_1^\dagger) + \log(2)$	Table 5b
	$\gamma_{int} > \log(2)/2$	Table 5c
G ₂ Monotonic	$\gamma_{int} > \log(3)$	Table 5a
	$\gamma_{int} > \log(3)/2$	Table 5c
No Assumption	$\gamma_{int} > \log(5)$	Table 5a
	$\gamma_{int} > \log(4)$	Table 5d
	$\gamma_{int} > \log(4)/2$	Table 5c

Table 8
Tests for Compositional Epistasis Between HLA-DRB1 SE and minor R620W PTPN22 alleles

Assumption	Condition on Model (4)	Form of Epistasis	Estimate of contrast And 95% CI	Satisfied by estimate	Satisfied by C.I.
G ₁ and G ₂ Monotonic	$\lambda_{int}/\mu > 0$	Table 5a	5.6 (3.3,7.9)	Yes	Yes
	$\lambda_{int}/\mu - (\alpha_2/\mu - \alpha_1/\mu) > 0$	Table 5b	-4.6 (-8.4, -0.8)	No	No
G ₁ Monotonic	$\lambda_{int}/\mu - \alpha_1/\mu > 0$	Table 5a	0.9 (-1.4,3.2)	Yes	No
	$\lambda_{int}/\mu - (\alpha_2/\mu - \alpha_1/\mu) - 1 > 0$	Table 5b	-5.6 (-9.4,-1.6)	No	No
	$\lambda_{int}/\mu - 1/2 > 0$	Table 5c	5.1 (2.8,7.4)	Yes	Yes
G ₂ Monotonic	$\lambda_{int}/\mu - \alpha_1/\mu - \beta_1/\mu - 2 > 0$	Table 5a	-0.7 (-3.1,1.8)	No	No
	$\lambda_{int}/\mu - (\alpha_1/\mu + 2) > 0$	Table 5d	-0.1 (-2.4,2.1)	No	No
	$\lambda_{int}/\mu - (\alpha_1/\mu + 2)/2 > 0$	Table 5c	4.8 (2.5,7.0)	Yes	Yes
No Assumption	$\lambda_{int}/\mu - 2\alpha_1/\mu - \beta_1/\mu - 4 > 0$	Table 5a	-6.4 (-9.4,-3.4)	No	No
	$\lambda_{int}/\mu - 2\alpha_1/\mu - 3 > 0$	Table 5d	-4.9 (-7.6,-2.1)	No	No
	$\lambda_{int}/\mu - (\alpha_1/\mu + 3)/2 > 0$	Table 5c	2.3 (0.0,4.5)	Yes	Yes