Invariants and noninvariants in the concept of interdependent effects

by Sander Greenland, DrPh, MS,[1] Charles Poole, MPH[2]

GREENLAND S, POOLE C. Invariants and noninvariants in the concept of interdependent effects. *Scand J Work Environ Health* 14 (1988) 125—129. In two of his publications [Causal and preventive interdependence: Elementary principles. *Scand J Work Environ Health* 8 (1982) 159—168 and *Theoretical Epidemiology,* John Wiley & Sons, New York, NY 1985], Miettinen put forth basic definitions of causal and preventive interdependence of effects involving binary exposure indicators and outcomes. This paper shows that the identification of interdependence using Miettinen's definitions varies with the choice of the reference categories for the exposures. In particular, Miettinen's concepts of synergism and antagonism are not invariant under exposure recoding. It is also shown that, when both exposures affect risk in some individuals, the effects will appear interdependent under some choice of referent. In the deterministic case, invariant properties of joint effects may be identified through the formation of equivalence classes of response types. In the stochastic case, invariant properties may be identified through the averaging of individual hazards, rather than risks. In both cases, additivity of risk or rate differences emerges as an elementary criterion for the independence of effects.

Key terms: additive model, effect modification, epidemiologic methods, interaction, synergy.

The definitions of synergy, antagonism, and biological interaction have been the matter of some controversy in epidemiologic literature (1, 3—14). Miettinen (5, 6) put forth a set of definitions of causal interdependence of effects based on individual disease outcomes under different exposure patterns as a solution to the controversy. The present paper shows that, under Miettinen's definitions, the interdependence of effects varies with changes in the exposure reference categories. Nevertheless, there are reference-invariant properties of joint effects that can be studied, even in the absence of "natural" reference categories.

Notation and definitions

To describe the basic concepts, suppose that we study two binary exposure variables, X and Y, each with possible values of 1 and 0. The object of study is the impact of changes in X and Y on the risk of a binary disease outcome D. For each individual i under study, let r_{xyi} be a binary indicator for the individual's disease outcome when the exposure levels are x and y (for example, $r_{10i} = 1$ if individual i gets the disease when $X = 1$ and $Y = 0$, but is zero otherwise). Finally, suppose 1 is considered the "exposed" (index value of

X and Y, and 0 is considered the "unexposed" (reference) value.

There are $2^4 = 16$ possible types of individual response patterns. Table 1 presents these 16 patterns. Types 8, 10, 12, and 14 were classified by Miettinen as instances of *causal interdependence* (interdependent

Table 1. Enumeration of possible response configurations to the four possible exposure combinations.

Type	Exposure combinations				Description [a]
	$X = 1$ $Y = 1$	$X = 0$ $Y = 1$	$X = 1$ $Y = 0$	$X = 0$ $Y = 0$	
1	1	1	1	1	No effects (doomed)
2	1	1	1	0	Single plus joint causation by $X = 1$ and $Y = 1$
3 [b]	1	1	0	1	$Y = 1$ blocks $X = 1$ effect (preventive antagonism)
4	1	1	0	0	$X = 1$ ineffective, $Y = 1$ causal
5 [b]	1	0	1	1	$X = 1$ blocks $Y = 1$ effect (preventive antagonism)
6	1	0	1	0	$X = 1$ causal, $Y = 1$ ineffective
7 [b]	1	0	0	1	Mutual blockage (preventive antagonism)
8 [b]	1	0	0	0	$X = 1$ plus $Y = 1$ causal (causal synergism)
9 [b]	0	1	1	1	$X = 1$ plus $Y = 1$ preventive (preventive synergism)
10 [b]	0	1	1	0	Mutual blockage (causal antagonism)
11	0	1	0	1	$X = 1$ preventive, $Y = 1$ ineffective
12 [b]	0	1	0	0	$X = 1$ blocks $Y = 1$ effect (causal antagonism)
13	0	0	1	1	$X = 1$ ineffective, $Y = 1$ preventive
14 [b]	0	0	1	0	$Y = 1$ blocks $X = 1$ effect (causal antagonism)
15	0	0	0	1	Single plus joint prevention by $X = 1$ and $Y = 1$
16	0	0	0	0	No effects (immune)

[a] Name according to Miettinen's classification in parentheses.
[b] Interdependence of effects, according to Miettinen (4).

[1] Division of Epidemiology, UCLA School of Public Health, Los Angeles, California, United States.
[2] Epidemiology Resources Inc, Chestnut Hill, Massachusetts 02167, and the Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, United States.

effects, with both exposures causal), and types 3, 5, 7, and 9 were classified as instances of *preventive interdependence*. For example, an individual of type 8, for whom disease occurs only if both exposures are present, was defined as a synergistic responder.

Impact of change of reference category

Consider now what happens to the classification scheme in table 1 when the index and reference categories of the second (Y) variable are interchanged. For example, suppose we excluded individuals of indeterminate sex from our universe of discourse, and suppose that Y was sex, $Y = 1$ indicated "male," and $Y = 0$ indicated "female." If we now redefine $Y = 1$ as "female," individuals coded $Y = 1$ before would now be recoded as $Y = 0$, and vice versa.

One consequence of the recoding of Y would be that individuals classified as synergistic (type 8) responders before would now be classified as antagonistic (type 14) responders. Thus the type of interdependence varies with the coding convention. Note also that type 15 responders, who were not classified as exhibiting interdependent effects in Miettinen's system, would become type 12 responders, whom he classified as exhibiting an antagonistic response pattern. Thus, when Miettinen's definitions are used, the classification of effects as interdependent can also change upon recoding.

Without the adoption of a deeper theory about how effects result, no reference category can be regarded as correct. For example, gender is a powerful determinant of risk for many diseases, yet there is often no basis for claiming that elevated risk among men represents causation of disease as a consequence of being male (eg, testosterone) rather than prevention of disease as a consequence of being female (eg, estrogen). In an analogous fashion, there will often be no basis for claiming that the transadditivity of risk differences seen when "female" is taken as the referent implies synergism between maleness and the other factor (cf equation 3 in reference 5) rather than antagonism between femaleness and the other factor.

Another implication of the effect of recoding is that it is ambiguous to say that two *variables* act synergistically in much the same way that it is meaningless to say a *variable* affects risk. The effects on risk are meaningfully defined in terms of the index and reference *levels* of a variable (6) or factors corresponding to the index levels.

In light of the preceding observations, it would seem worthwhile to discover what, if any, properties of joint action are invariant under recoding or changes in reference levels.

Invariant properties of response types

To find the invariant properties of the response types in table 1, we will partition the types into *equivalence*

classes based on recoding of the exposure indicators. Two response types (m and n) will be defined as *equivalent under recoding* (or simply equivalent) if there exists a one-to-one recoding of X and/or Y such that every individual of type m becomes an individual of type n upon recoding. For example, type 8 and type 14 are equivalent because, upon recoding y as $1 - y$ (which corresponds to interchanging the index and reference categories for Y), type 8 individuals become type 14 individuals. Likewise, type 8 and type 15 are equivalent because, upon recoding x as $1 - x$ and y as $1 - y$, type 8 individuals become type 15 individuals.

It is an easy algebraic exercise to verify that the aforementioned relation is a formal equivalence relation: it is (i) reflexive, ie, type m is equivalent to type m; (ii) symmetric, ie, if type m is equivalent to type n, then type n is equivalent to type m; and (iii) transitive, ie, if type m is equivalent to type n and type n is equivalent to type p, then type m is equivalent to type p. Thus we can form mutually exclusive and exhaustive equivalence classes of types, the members of each class being equivalent to one another but not to members of other classes. There are seven such classes, which we label as follows: (i) C_D, comprising only type 1, the doomed type; (ii) C_X, comprising types 6 and 11, ie, both the types in which $X = 1$ is effective but $Y = 1$ is not; (iii) C_Y, comprising types 4 and 13, ie, both the types in which $Y = 1$ is effective but $X = 1$ is not; (iv) C_M, comprising types 7 and 10, ie, both the types exhibiting mutual antagonism; (v) C_S, comprising types 8, 12, 14, and 15, ie, all the types in which disease occurs for only one exposure combination; (vi) C_T, comprising types 2, 3, 5, and 9, ie, all the types in which disease occurs for three exposure combinations; (vii) C_I, comprising only type 16, the immune type. As we should hope, lack of an effect of either or both factors (classes C_D, C_X, C_Y, and C_I) turns out to be invariant under recoding. In contrast, of the earlier classification of types of interdependent action, only mutual antagonism (class C_M) remains invariant under recoding. Causal synergism (type 8), asymmetric causal antagonism (types 12 and 14), and single-plus-joint prevention (type 15) turn out to be equivalent to one another (class C_S). Similarly, single-plus-joint causation (type 2), asymmetric preventive antagonism (types 3 and 5), and preventive synergism (type 9) turn out to be equivalent (class C_T). Thus neither synergism, asymmetric antagonism, nor single-plus-joint action are invariant properties.

At first sight, the equivalence (within classes C_S and C_T) of two types previously classified as exhibiting interdependent action (synergism and asymmetric antagonism) with a type previously classified as exhibiting independent action (single-plus-joint action) may seem anomalous. However, type 2 can be viewed as synergistic prevention by $X = 0$ and $Y = 0$, and type 15 can be viewed as synergistic causation by $X = 0$ and $Y = 0$. Thus failure to recognize types 2 and 15 as exhibiting interdependent action can be seen as a con-

sequence of choosing $X = 1$ and $Y = 1$ as the "exposed" categories, and nothing more. If this choice is arbitrary, types 2 and 15 represent interdependence as much as the other types in classes C_S and C_T. All types in class C_S (and no other types) can be viewed as representing synergistic causation under recoding, and all types in class C_T (and no other types) can be viewed as representing synergistic prevention under recoding.

Implications for incidence

Consider now a fixed population of individuals, with a proportion (P_m) of individuals of type m ($\Sigma_m P_m = 1$). Let R_{xy} be the proportion which becomes ill when $X = x$ and $Y = y$. Miettinen (5) described the mathematical relations among the differences $R_{xy} - R_{00} = RD_{xy}$ implied by the absence from the population of various response types. In an analogous fashion, we will present mathematical relations among the R_{xy} and P_m implied by the absence of certain equivalence classes of types. It will be convenient to employ Koopman's "interaction contrast" (3), which is defined as

$$I = RD_{11} - (RD_{10} - RD_{01}) = R_{11} - R_{10} - R_{01} + R_{00}.$$
$$= P_3 + P_5 + 2P_7 + P_8 + P_{15} - P_2 - P_9 - 2P_{10} - P_{12} - P_{14}.$$

Suppose first that none of the types in classes representing causal interdependence (C_M, C_S, and C_T) are present. This supposition is equivalent to assuming the presence of types 1, 4, 6, 11, 13, and 16 only. Then $I = 0$, so that $RD_{11} = RD_{10} + RD_{01}$. Miettinen (5) found the same additivity relation assuming no causal interdependence, purely causal effects, and "perfect negative correlation of susceptibilities." These conditions are equivalent to assuming the absence of all types but 1, 4, 6, and 16. Our condition is slightly more general since it allows mixtures of causal and preventive actions. More importantly, we have shown that this relationship is invariant under the arbitrary recoding of X and Y (ie, arbitrary exchange of reference and index levels of the variables). Thus, under the deterministic model, departures from the additivity of risk differences always imply the presence of types in one of the equivalence classes representing interdependent action, ie, C_M, C_S, or C_T.

Type 2 is usually considered a possible response pattern, and so it is worth considering the situation in which types in its class, C_T, are present. Suppose then that only classes C_M and C_S are absent, ie, that types 7, 8, 10, 12, 14, and 15 are absent. Then $I = P_3 + P_5 - P_2 - P_9$, so that I is a contrast of proportions of all types in class C_T. Thus departures from additivity (ie, $I \neq 0$) imply the existence of C_T types in the population (although not vice versa, since $I = 0$ if $P_3 + P_5 = P_2 + P_9$). Again, this is a coding-invariant rule. Under the additional (noninvariant) assumption of no prevention, we find that $\min(RD_{10} +$

$RD_{01}, 1 - R_{00}) \geq RD_{11} \geq \max(RD_{10}, RD_{01})$ as shown by Miettinen (equation 1 in reference 5).

In an analogous fashion, if classes C_M and C_T are absent, departure from additivity invariantly implies the existence of C_S types. The additional noninvariant assumption of no causation implies that $\min(RD_{10}, RD_{01}) \geq RD_{11} \geq \max(RD_{10} + RD_{01}, -RD_{00})$ (equation 2 in reference 5). Similarly, if classes C_S and C_T are absent, departure from additivity invariantly implies the existence of C_M (mutual antagonism) types.
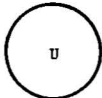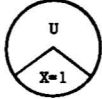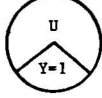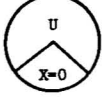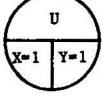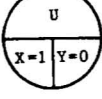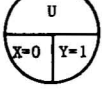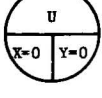
The stochastic case
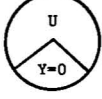
For the stochastic case, r_{xyi} is now the individual probability (risk) of disease at exposure level x, y; in other words, r_{xyi} may now take on values between zero and one. In what follows, the subscript i will, for simplicity, be omitted. Miettinen (5) reasoned that the independence of causal effects in this case corresponded to probabilistic independence of effects, ie, $r_{11} = r_{10} + r_{01} - r_{10}r_{01}$. Letting $s_{xy} = 1 - r_{xy}$, this condition may be rewritten as $s_{11} = s_{10}s_{01}$. With the assumption that s_{xy} is strictly positive, we define the cumulative hazard $h_{xy} = -\log_e s_{xy}$. Then the independence condition may be rewritten as $h_{11} = h_{10} + h_{01}$. This condition is *not* invariant under recoding, since it imposes no strict constraints on the relation of h_{00} to the other hazards.

Consider now a situation in which the effects of the exposures ($X = 1$ and $Y = 1$) are probabilistically independent of any background causes, as well as of one another's effect, so that $s_{10} = s_X s_0$, $s_{01} = s_Y s_0$, and $s_{11} = s_X s_Y s_0$, where s_X, s_Y, and s_0 are the individual's probability of escaping disease produced by X exposure, Y exposure, and other causes, respectively. This situation is equivalent to one in which the cumulative hazards follow the additive model $h_{xy} = h_0 + h_X x + h_Y y$, where $-h_X$, $-h_Y$, and $-h_0$ are the natural logarithms of s_X, s_Y, and s_0. It is easily verified that this additive structure is invariant under recoding (although the value of h_0 and the signs of h_X and h_Y will change).

Note that the aforementioned model can be rewritten to incorporate preventive as well as causal action. For example, if both $X = 1$ and $Y = 1$ are preventive, independent action may be defined as $s_{00} = s_X s_Y s_1$, where s_X, s_Y, and s_1 are the probabilities of escaping disease produced by causes blocked by $X = 1$, causes blocked by $Y = 1$, and other causes, respectively. Here s_X and s_Y would be interpreted as causal probabilities if X and Y were recoded but would not change their numerical values.

Consider now a population of size N in which H_{xy}, H_X, H_Y, and H_0 are the mean values of the h_{xy}, h_X, h_Y, and h_0. The incidence proportions R_{xy} will approximate the H_{xy} if the r_{xy} are small, since in that case $r_{xy} \doteq h_{xy}$, and so $R_{xy} = \Sigma r_{xy}/N \doteq \Sigma h_{xy}/N = H_{xy}$,

127

where the sums are over individuals. It is apparent from the last equality in the preceding series that if h_{xy} follows the aforementioned additive model, then $R_{xy} = H_{xy} = H_0 + H_X x + H_Y y$. Thus for low-risk diseases the incidence proportions will approximately follow an additive model (with coefficients H_0, H_X and H_Y) if, for all individuals, the study factors and background factors act independently of one another; this property is invariant under recoding. Equation 6 of Miettinen (5) also showed that an approximate additive model holds if the mean of the $r_{10}r_{01}$ is negligible relative to $R_{10} + R_{01}$. We have shown that this result is invariant provided that all the r_{10} and r_{01} are also small (so that $R_{xy} \doteq H_{xy}$), and that the exposures act independently of background causes.

SUFFICIENT CAUSE TYPE	DESCRIPTION
1	X and Y irrelevant
2	X = 1 necessary, Y irrelevant
3	Y = 1 necessary, X irrelevant
4	X = 0 necessary, Y irrelevant
5	Y = 0 necessary, X irrelevant
6	X = 1 and Y = 1 necessary
7	X = 1 and Y = 0 necessary
8	X = 0 and Y = 1 necessary
9	X = 0 and Y = 0 necessary

U = all other components of the sufficient cause

Figure 1. Enumeration of the nine types of sufficient causes for two dichotomous exposure variables.

When the distribution of disease occurrence time is of primary interest (as for most chronic diseases of adults), the individual probabilities and hazards become functions of time, and the independent-action model becomes $s_{xy}(t) = s_X(t)^x s_Y(t)^y s_0(t)$. This result implies $h_{xy}(t) = h_0(t) + h_X(t)x + h_Y(t)y$, where the h's now represent instantaneous hazards [ie, the derivative of the cumulative hazard $-\log_e s(t)$ at t]. The population mean $H_{xy}(t)$ of $h_{xy}(t)$ at t is just the instantaneous incidence density at level x, y (2). Under the independent-action model, $H_{xy}(t) = H_0(t) + H_X(t)x + H_Y(t)y$, where $H_0(t)$, $H_X(t)$, and $H_Y(t)$ are means of $h_0(t)$, $h_X(t)$, and $h_Y(t)$. Again, this additivity of the incidence densities is invariant under recoding. Note however that the coefficients in this model, $H_0(t)$, $H_X(t)$, and $H_Y(t)$, may vary independently of one another over time. Thus the model is much more general than commonly used additive models, such as $H_{xy}(t) = (1 + b_1 x + b_2 y)H_0(t)$, in which the coefficients maintain a fixed proportion to one another over time.

Unlike the results given by Miettinen (5), the preceding results employ the assumption that the actions of the exposures are independent of background factors. With or without this assumption, the hypothesis that the exposure effects are independent would be refuted in principle if RD_{11} does not fall between $\max(RD_{10}, RD_{10})$ and $RD_{10} + RD_{01}$ (5). With this assumption, however, the hypothesis of independent effects would be refuted in principle by any departure of the instantaneous incidence densities from additivity.

Correspondence to the sufficient-component cause model

The discussion thus far has concerned only the classification of biological processes according to their manifestations in individual risk patterns. A deeper theory would begin with a classification of biological processes from which these manifestations would be derived. The simplest of these theories is perhaps the sufficient-component cause theory described by Rothman (7, 8). This deterministic biological theory has some interesting relations to the equivalence classes given earlier for the deterministic case.

Figure 1 presents the nine types of sufficient causes possible under the sufficient-component cause theory. Note that, whereas response types refer to individuals, the cause types in figure 1 refer to causes. We will present a correspondence between sufficient cause types and individual response types.

We will say a person is "at risk" for (or susceptible to) a particular type of cause if that person has all the components of a sufficient cause of that type with the possible exception of the components (if any) from X, Y, or both. Note that a person is of one and only one response type, whereas a person can be at risk of more

than one type of sufficient cause. Note also that, by the definition of sufficient cause, only one sufficient cause can be responsible for a particular occurrence of disease.

We can form equivalence classes of these sufficient-cause types by examining what types are mapped onto one another upon the recoding of X, Y, or both. Upon doing so, we find that there are four such classes: (i) U_0, comprising only cause type 1, in which neither X nor Y contribute a component; (ii) U_X, comprising types 2 and 4, in which X but not Y contributes a component; (iii) U_Y, comprising types 3 and 5, in which Y but not X contributes a component; (iv) U_{XY}, comprising types 6—9, in which both X and Y contribute a component. The invariant properties are simple, ie, no involvement of either variable, involvement of only X, involvement of only Y, and involvement of both. Sufficient causes in classes U_X and U_Y represent independent action of X and Y, and causes in class U_{XY} represent coaction of X and Y (coparticipation in a sufficient cause).

Given a list of the cause types for which a person is at risk, we can deduce the person's response type. For example, a person at risk of cause type 1 will be of response type 1, for that person will get the disease regardless of that person's X or Y status; a person at risk of cause types 2, 3, and no others will be of response type 2; and a person at risk of cause types 6, 7, 8, and no others will also be of response type 2.

The last two examples illustrate the fact that most response types may arise from more than one set of cause types. In this sense, the set of cause types for which a person is at risk is not always identifiable given the person's response type. Nevertheless, there are several important exceptions to this rule. Each response type in class C_S (types 8, 12, 14, 15) is produced only by being at risk of exactly one cause type from class U_{XY} (types 6, 7, 8, 9). For example, a person is of response type 15 if, and only if, that person is at risk of cause type 6 and no others. Thus C_S, the response class corresponding to synergistic causation, has a one-to-one correspondence with U_{XY}, the sufficient-cause class corresponding to causal coaction.

A person is of response type 16 (no response under any conditions) if that person is not at risk of any of the cause types in figure 1. Unfortunately, no other one-to-one correspondences arise for the remaining 11 response types.

Discussion

Miettinen (5, 6) constructed his approach on the assumption that causal and preventive action could be sharply distinguished as ontological concepts. In contrast, the present approach is based on a more general and primitive ontology in which causal and preventive actions are distinguished only by arbitrary coding. It should be noted that the choice between the two ontologies is not a matter of correctness, but rather one of the degree of specification appropriate for the existing state of knowledge. Certainly, a situation as well elaborated as measles control, for example, calls for a sharp ontological distinction between the preventive action of a vaccine and the causal action of the virus. Such a degree of elaboration is certainly the preferable one. But, as illustrated earlier for gender effects, epidemiologists are often confronted with effects for which it is far from clear whether the mechanism of action should be classified as preventive or causal (and thus for which it is also far from clear what the appropriate reference and index categories should be). In such situations it remains meaningful to ask whether the effects of various factors are independent. As previously demonstrated, partial answers to this question may be obtained with estimates of rates or average risks (just as in the more elaborated situation considered by Miettinen), even if questions concerning synergy and antagonism must be put aside.

It is interesting to note that, if the definitions and the simplest causal models given in this paper are used, additivity of risk or rate differences again emerges as the most basic manifestation of independent effects. Nevertheless, at least in the stochastic case, other concepts of independent effects may be possible, and these would have different manifestations.

References

1. Blot WJ, Day NE. Synergism and interaction: Are they equivalent? Am J Epidemiol 110 (1979) 99—100. (Letter to the editor).
2. Greenland S. Interpretation and choice of effect measures in epidemiologic analyses. Am J Epidemiol 125 (1987) 761—768.
3. Koopman JS. Interaction between discrete causes. Am J Epidemiol 113 (1981) 716—724.
4. Kupper LL, Hogan MD. Interaction in epidemiologic studies. Am J Epidemiol 106 (1978) 447—453.
5. Miettinen OS. Causal and preventive interdependence: Elementary principles. Scand J Work Environ Health 8 (1982) 159—168.
6. Miettinen OS. Theoretical epidemiology. John Wiley & Sons, New York, NY 1985.
7. Rothman KJ. Causes. Am J Epidemiol 104 (1976) 587—592.
8. Rothman KJ. Modern epidemiology. Little, Brown, Boston MA 1986.
9. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. Am J Epidemiol 112 (1980) 467—470.
10. Saracci R. Interaction and synergism. Am J Epidemiol 112 (1980) 465—466.
11. Siemiatycki J, Thomas DC. Biological models and statistical interactions: An example from multistage carcinogenesis. Int J Epidemiol 10 (1981) 383—387.
12. Vandenbroucke JP. More on independence and interaction. Am J Epidemiol 114 (1981) 167. (Letter to the editor).
13. Walter SD, Holford TR. Additive, multiplicative, and other models for disease risks. Am J Epidemiol 108 (1978) 341—346.
14. Weed DL, Selmon M, Sinks T. Links between categories of interaction. Am J Epidemiol 127 (1988) 17—27.