Scand J Work Environ Health 1982;8(3):159-168

<https://doi.org/10.5271/sjweh.2479>

Issue date: Sep 1982

Causal and preventive interdependence. Elementary principles.

by [Miettinen OS](#)

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/7156935

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Causal and preventive interdependence

Elementary principles

by Olli S Miettinen, MD, PhD¹

MIETTINEN OS. Causal and preventive interdependence: Elementary principles. *Scand j work environ health* 8 (1982) 159—168. "Synergism" of two factors in the causation or prevention of an all-or-none event means the existence of instances in which both factors are needed for the effect, while "antagonism" means that at least one can block the solo effect of the other. The manifestation of such interdependences — or of their complement, independence — in terms of event rates is complicated by the correlation of the susceptibilities to the two factors. Thus, given the risk difference (RD) values representing the solo effects, the RD corresponding to the joint exposure has a range consistent with independence so that independence cannot be inferred even from very ample data without knowledge of the degree of correlatedness of the susceptibilities. The definition of this range is closely analogous for causal and preventive factors, respectively. However, when knowledge about interdependence is used in inference about the factors' interrelation in the mechanisms for the effect, sharp distinctions may have to be made between causal and preventive factors. In each case, the interdependence is a result of the interrelation of the actions of the two factors and/or interaction between them. Operational decisions, having to do with the wisdom of the joint exposure, can be guided by knowledge about the interdependence of the factors; however, knowledge of the risks corresponding to the various exposures is a sufficient guide, without any need for inferences about causal or preventive interdependence.

Key terms: biometry, epidemiologic methods, statistics.

Even though causal and preventive interdependences — synergism and antagonism — are of great interest to epidemiologists, the discussion of their formal aspects is still being distracted by the continuing tendency of some to confuse these substantive matters of science with the arbitrary choice to use a nonadditive expression for the model which relates the outcome parameter to the factors at issue. One consequence of this distraction may be that even those epidemiologists who understand the distinction (12, 13) have had very few concrete and tenable things to say about even the most elementary issues regarding causal or preventive interdependence.

Thus Rothman et al (12) recently "defined" causal interdependence (they termed it "interaction") simply, and almost tautologically, as the "interdependent operation" of causes. A more-detailed definition they said, expressly, would be unnecessary, and they emphasized instead the importance of an explicit model. As for the latter, they singled out only the interchangeability of the causes in the same step of the causal process and their action in different steps, and they opined that these correspond generally to additive and multiplicative models, respectively.

While some already consider the topic "terribly overworked" (2), I, like many others, continue to be concerned with the clarity of the very first principles, both ontological and epistemological, of causal and preventive interdependence. It is my intention in the present paper (i) to attempt to improve the conceptualizations of the constituent entities of synergism and an-

¹ Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston, Massachusetts, United States.

Reprint requests to: Prof OS Miettinen, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA.

tagonism, (ii) to delineate the manifestations of these entities in epidemiologic data, and (iii) to touch upon the implications of having inferred either synergism or antagonism of causal or preventive factors. I have confined my analysis to the case of all-or-none effects.

Definitions

The definition of "synergy" in a Webster dictionary (18) is "combined action or operation (as of muscles)," while "synergism" is defined as "cooperative action of discrete agencies such that the total effect is greater than the sum of the effects taken independently." Dorland's medical dictionary (17) defines "synergy" as "correlated action or cooperation..." and "synergism" as "the joint action of agents so that their combined effect is greater than the algebraic sum of their individual effects."

Evidently, synergism has to do with *cooperation* (rather than interaction). In some sense, more is accomplished together than separately. A bit more specifically, the joint effect is said to exceed the sum of the separate effects (17, 18). But since the scale on which the effect is to be expressed is left unspecified, the concept remains vague (just as it is with the definition of Rothman et al as a basis — as they themselves point out).

For "antagonism" the Webster (18) definition is "opposition in physiological action; esp: interaction of two or more substances such that the action of any one of them on living cells or tissue is lessened." Dorland (17) gives "opposition or contrariety, as between muscles, medicines, or organisms."

The case of deterministic effects

In seeking to develop more explicit definitions for causal and preventive "synergism," "antagonism," and "independence" (in the context of events of the all-or-none sort), I shall first focus on *deterministic* effects, ie, factors which in any given instance either necessarily produce the effect or necessarily fail to produce it.

Example 1. Consider the responses of each of a set of five subjects (people) to two deter-

ministically causal exposures, E_1 and E_2 , both separately and jointly. Suppose the numbers of "reactors" are one with E_1 alone, one with E_2 alone, and two with E_1 and E_2 together. The solo effect for each exposure is one (out of five), and the joint effect two (out of five). The preceding definitions could be taken to mean absence of synergism in this experience — With no need for further specifications, the joint effect is the sum of the solo effects (on the scale of proportions).

Example 2. Consider potential details for example 1. Suppose first that E_1 alone is effective for subject 1, E_2 alone for subject 2, and E_1 and E_2 for subjects 1 and 2. Lack of synergism could be concluded as in example 1. But suppose now that both solo exposures are effective for subject 1, while $E_1 + E_2$ is effective for subjects 1 and 2. The preceding definitions would still mean absence of synergism. Yet the fact is that, in the latter case, the effectiveness of $E_1 + E_2$ in subject 2 is singularly the result of cooperation between the two exposures; also, for this subject the joint effect of the two exposures is larger than the sum of the solo effects (which were nil). Thus, subject 2 exemplifies synergism in the very sense of the Webster and Dorland definitions, with no concern as to the scale on which the effects are expressed.

Example 3. Suppose now that the details of the pattern of responsiveness in example 1 are as follows: E_1 and E_2 applied singly are effective for subjects 1 and 2, respectively, while $E_1 + E_2$ is effective for subjects 3 and 4. In this case, subjects 3 and 4 represent instances of synergism (cf example 2), while subjects 1 and 2 are cases of antagonism. For subject 1, E_2 counteracts E_1 , while the converse is the case for subject 2.

Example 4. Examples 1–3 deal with causation. Without exposure each of the subjects would have failed to experience the event, while the exposures had the potential of bringing it about, ie, the potential of changing absence of event to presence of it (— to +). These examples can be translated to ones of *prevention* with the supposition that in the absence of exposure each subject would experience the event, while exposure has the potential of changing the outcome from event to lack of it (+ to —).

From examples 1–4 it is evident that the definitions of Webster and Dorland are too superficial to be satisfactory. These examples also underscore the need to consider the existence of *instances* of certain patterns of responsiveness (no solo effects but response to the joint exposure, etc) as the basis for describing two exposures as synergistic and/or antagonistic — just as the existence of instances of effect is the basis of characterizing an ex-

posure as a cause and/or preventive (cf example 4). The term "instance," as used here, does not refer to some particular subject per se, because a given subject's response can depend on its/his/her environment and/or behavior; it refers to some particular subject behaving in a particular way in a particular environment. Thus even a small coronary atheroma ("exposure") is a cause of angina pectoris insofar as there are instances of subject/behavior/environment in which it brings about the symptom, eg, untrained persons (subjects) exerting themselves (behavior) in cold weather (environment).

The presented examples suggest that an *instance of synergism* between two deterministically causal (or preventive) factors in the context of an all-or-none event (illness) is one in which both exposures are needed to bring about the effect. In these instances each exposure is a sufficient cause (or preventive) in the presence of the other, ie, the two exposures jointly constitute a sufficient cause (or preventive), while each of them alone is insufficient; the phenomenon is symmetrical between the two factors involved.

An *instance of antagonism* between two deterministic factors is one in which at least one of the two is effective when alone (as a cause or as a preventive), but the combination is devoid of effect. In other words, at least one of the factors neutralizes (blocks) the effect of the other. Thus instances of antagonism are of three kinds. In mutual antagonism each of the factors, say E_1 and E_2 , blocks the effect of the other; in other instances E_1 may an-

tagonize E_2 , while E_2 does not antagonize E_1 (because the effect of E_1 alone is nil); finally, there may be instances of E_2 antagonizing E_1 without E_1 antagonizing E_2 (because the effect of E_2 alone is nil). These definitions are summarized in table 1.

The presented concepts may be further sharpened by a consideration of the relation between instances of prevention and those of antagonism. A preventive frustrates a sufficient background cause and is devoid of effect in the absence of such a background (cf example 4). Thus a preventive antagonizes background causes, while the converse is not the case, ie, instances of prevention can be thought of as occurrences of exposure-on-background antagonism.

"*Synergism*" of two deterministically causal (or preventive) factors — in bringing about (or preventing) a particular all-or-none outcome — may now be defined as the existence of instances in which both factors are needed in order to produce the effect, while "*antagonism*" refers to the existence of instances in which (at least) one blocks the effect of the other.

Examples 1–4 illustrate that in any given domain of instances (of particular subjects behaving in particular ways in their particular environments) there may be instances of synergism without any instances of antagonism, instances of antagonism without any instances of synergism, both types of instances, or neither. With reference to such domains, the two factors may be said to have *simple synergism*, *simple antagonism*, *mixed interde-*

Table 1. Instances^a of synergism and antagonism defined for two deterministically causal or preventive exposures, E_1 and E_2 , according to outcomes (+ or —) associated with absence of both exposures ($\bar{E}_1\bar{E}_2$), presence of E_1 alone ($E_1\bar{E}_2$), etc.

Type of instance ^a	Definitional pattern of outcome								Prevalence of type
	A. Exposures causal				B. Exposures preventive				
	$\bar{E}_1\bar{E}_2$	$E_1\bar{E}_2$	$\bar{E}_1E_2$	E_1E_2	$\bar{E}_1\bar{E}_2$	$E_1\bar{E}_2$	$\bar{E}_1E_2$	E_1E_2	
Instance of synergism (S)	—	—	—	+	+	+	+	—	S
Instance of antagonism (A)	—	(+) ^b	—	—	+	(—) ^b	—	+	A
Mutual	—	+	+	—	+	—	—	+	A _{ab}
E ₁ on E ₂	—	—	+	—	+	+	—	+	A _a
E ₂ on E ₁	—	+	—	—	+	—	+	+	A _b

^a Instance = a particular subject behaving in a particular way in a particular environment.

^b E_1 and/or E_2 present alone has this effect.

pendence (synergism and antagonism), and (simple) *independence*, respectively. Where both synergism and antagonism occur, depending on the instance, the balance is one of synergism, antagonism, or independence. With the population fractions representing instances of synergism and antagonism denoted by S and A (cf table 1), respectively, these definitions imply (i) simple synergism if $S > 0$, $A = 0$; (ii) simple antagonism if $S = 0$, $A > 0$; (iii) synergism and antagonism if $S > 0$, $A > 0$ with synergism in balance if $S > A$, antagonism in balance if $S < A$, independence in balance if $S = A$; and (iv) simple independence if $S = 0$, $A = 0$.

The general case

The issues are a bit more subtle in the context of stochastic causation (or prevention) — where the effect in any given instance is a change phenomenon, subject only to a probabilistic characterization. The change aspect then refers expressly to the way in which the causes operate rather than to uncertainty about the particular of the instances in which they operate, because the concern is with ontological formulations rather than with epistemological problems. Stochastic causation (or prevention) in this sense is, of course, a generalization of deterministic effects, ie, the latter is a special case of the former.

An *instance of synergism* of two factors in the context of stochastic causation (or prevention) — as a generalization of the definition in the context of deterministic effects — is one in which a factor, when failing to bring about the effect on its own, enhances the probability that the other factor will produce it. (In the context of deterministic effects the probability changes from zero to one.)

Analogously, an *instance of antagonism* is one in which the presence of one factor, when it is itself without effect, reduces the probability that the other factor will have the effect.

As before, “*synergism*” and “*antagonism*” refer to the existence of the respective instances, and (simple) “*independence*” to their absence. Where both occur (mixed interdependence), the overall pattern (synergism, antagonism, or independence)

is determined by the balance of the positive and negative changes of probability corresponding to the individual instances of synergism and antagonism, respectively.

Example 5. Suppose each member of a particular domain of subjects represents an instance of not experiencing the event at issue in the absence of both of two exposures and suppose also that each of the two exposures alone makes the probability of the event equal to $1/2$ for each subject. In the case of simple independence between the two effects in the domain at issue the joint exposure makes the probability equal to $1/2 + (1 - 1/2)1/2 = 3/4$ for each subject. Simple synergism means that with joint exposure the probability exceeds $3/4$ for at least one subject and it falls short of $3/4$ for none of them. In simple antagonism it is less than $3/4$ for at least one subject and greater than $3/4$ for none of them. Mixed dependence occurs when there are instances of joint exposure for which the probability of the event exceeds $3/4$ and others for which it is less than $3/4$. If with mixed dependence the average probability exceeds $3/4$, there is synergism in balance; if it falls short of $3/4$, the overall pattern is one of antagonism; and, if the average probability is $3/4$, then the mixed interdependence amounts to independence in balance.

Epidemiologic manifestations

The detection of synergism and/or antagonism may be based directly on the definitions already presented if it is possible to ascertain *individual susceptibilities* (causal or preventive) to each of the two exposures separately and also to joint exposure to the two. (This ascertainment requires reasonably acute and reversible effects and feasibility of individual cross-over trials.) Otherwise — and this is generally the case in epidemiology — it is necessary to make do with comparisons of the *rates of occurrence* of the outcome (illness) of interest among the four domains of exposure ($\bar{E}_1\bar{E}_2$, $E_1\bar{E}_2$, $\bar{E}_1E_2$, and E_1E_2 ; see table 1). Thus, for the purpose of the epidemiologic detection of synergism or antagonism, it is necessary to understand the manifestations of independence, synergism, and antagonism in terms of the comparative rates.

The case of deterministic effects

A couple of examples should serve to introduce the issues in the context of deterministic effects.

Example 6. Consider a set of five instances (of particular persons behaving in particular ways in particular environments), in which instances 1 and 2 represent susceptibility to deterministic exposure E_1 and instances 1 to 3 represent susceptibility to deterministic exposure E_2 (see table 2, panel A). In the absence of both synergism and antagonism, subjects 1, 2, and 3 are susceptible to the combined exposure, while subjects 4 and 5 are not. Thus the rate difference for the joint exposure (relative to neither exposure) in this example is the maximum of the rate differences for the two solo exposures, respectively (table 2, panel A).

Example 7. Suppose now that, while subjects 1 and 2 are susceptible to the first exposure, subjects 3, 4, and 5 are susceptible to the other exposure. In this situation, in the absence of both synergism and antagonism, all five subjects are susceptible to the joint exposure so that the rate difference for the joint exposure is the sum of those for the component exposures acting alone (table 2, panel B).

These examples show that, in the context of deterministic factors, the manifestation of causal (or preventive) independence in comparative rates depends on the *correlatedness of the susceptibilities* to the two factors (3, 14). In one extreme those susceptible to one factor are a subset of those susceptible to the other, and in this situation, of maximal correlation, the rate difference for the joint exposure (RD_{11}) is the maximum of the rate differences for those exposed to the first (RD_{10}) and second (RD_{01}) factor, respectively. In the other extreme the two domains of susceptibility are maximally disjoint, and in this case, of minimal (extreme

negative) correlation, RD_{11} may equal $RD_{10} + RD_{01}$ (cf example 7). However, RD_{11} is bound by $1 - R_{00}$ in the case of causation and by $-R_{00}$ in the context of preventive effects, where R_{00} is the proportion of subjects experiencing the event in the absence of both exposures. Thus *causal independence* means that the joint effect is bound as follows:

$$\max(RD_{10}, RD_{01}) \leq RD_{11} \leq \min(RD_{10} + RD_{01}, 1 - R_{00}); \quad (\text{eq 1})$$

and in the context of *preventive independence*, with $RD_{10} \leq 0$ and $RD_{01} \leq 0$, the range for RD_{11} is:

$$\max(RD_{10} + RD_{01}, -R_{00}) \leq RD_{11} \leq \min(RD_{10}, RD_{01}). \quad (\text{eq 2})$$

The argument leading to these relations was based on the assumption of *simple independence*. The end results do apply, however, to balanced dependence (independence-in-balance) as well: for the effect of solo exposure the distinction is totally moot by definition, while the joint exposure blocks the effect of solo exposure in some instances but has a unique effect of its own in a compensatory number of other instances; thus all three rate differences remain unchanged.

When the preceding relations (equations 1 and 2) are satisfied, the experience, while consistent with independence, does not expressly imply independence; they would

Table 2. Manifestation of the independent effects of deterministically causal factors in terms of rate differences (RDs),^a according to the correlatedness of susceptibilities and individual outcomes (+ or —).

Instance	A. Extreme positive correlation of susceptibilities				B. Extreme negative correlation of susceptibilities			
	$\bar{E}_1\bar{E}_2$	$E_1\bar{E}_2$	$\bar{E}_1E_2$	E_1E_2	$\bar{E}_1\bar{E}_2$	$E_1\bar{E}_2$	$\bar{E}_1E_2$	E_1E_2
1	—	+	+	+	—	+	—	+
2	—	+	+	+	—	+	—	+
3	—	—	+	+	—	—	+	+
4	—	—	—	—	—	—	+	+
5	—	—	—	—	—	—	+	+
RD	—	2/5 (RD_{10})	3/5 (RD_{01})	3/5 (RD_{11})	—	2/5 (RD_{10})	3/5 (RD_{01})	5/5 (RD_{11})
	$RD_{11} = \max(RD_{10}, RD_{01})$				$RD_{11} = RD_{10} + RD_{01}$			

^a RD = rate in the index domain (say, $E_1\bar{E}_2$) minus rate in the reference domain ($\bar{E}_1\bar{E}_2$), eg, RD_{10} is the rate in $E_1\bar{E}_2$ minus the rate in $\bar{E}_1\bar{E}_2$.

be satisfied in the context of some degree of synergism or antagonism as well. Thus independence cannot be inferred from event rates. On the other hand the rate difference for the joint exposure falling outside the unconstrained range does imply synergism or antagonism (apart from statistical uncertainty) — while the converse does not obtain:

$$RD_{11} > RD_{10} + RD_{01} \Rightarrow \begin{matrix} \text{synergism of} \\ \text{causes;} \end{matrix} \quad (eq\ 3)$$

$$RD_{11} < \max(RD_{10}, RD_{01}) \Rightarrow \begin{matrix} \text{antagonism} \\ \text{of causes;} \end{matrix}$$

$$RD_{11} < RD_{10} + RD_{01} \Rightarrow \begin{matrix} \text{synergism of} \\ \text{preventives;} \end{matrix} \quad (eq\ 4)$$

$$RD_{11} > \min(RD_{10}, RD_{01}) \Rightarrow \begin{matrix} \text{antagonism} \\ \text{of preventives.} \end{matrix}$$

The derivation of these criteria in the preceding equations indicates that they apply to rates in the sense of the *proportion* developing the event (illness) over the risk period of interest, ie, to cumulative incidence rates and risks over that period. When chronic etiology is concerned, the primary information may take the form of age-specific (or other interval-specific) incidence densities for each of the exposure domains in question, or, given a case-referent design, it can be translated into these (7). The above criteria can be applied to this case by translating the densities into cumulative incidence rates or risks (7) over the risk period of interest. Where the risks are substantial, the criteria for evident synergism (expressions 3 & 4), based on risk differences, do not apply to incidence-density differences.

Whereas in the simple, idealized case the same set of subjects would be studied under each of the four relevant types of exposure (see table 1), in ordinary epidemiologic practice the experiences are disjoint: there is one set of instances with both factors absent, another with the first factor present alone, etc. In this situation it is, of course, essential that the four rates, or the three rate differences, be mutually comparable. As for the proper meaning of "comparability" in the assessment of independence of effects, it is to be understood that the concern is not only with *confounders* (8) but with *modifiers* (6) — of unconfounded rate difference — as well.

The general case

The general case of stochastic causation (or prevention) brings nothing really new.

Consider causation first, along with a situation in which there are no sufficient "background" causes, ie, no "spontaneous" events.

Let the two exposures have solo effects with risks (probabilities) $R_{10,i}$ and $R_{01,i}$, respectively, in the i^{th} instance. If they are independent, as already defined, the probability of effect for joint exposure is:

$$R_{11,i} = R_{10,i} + R_{01,i} - R_{10,i} R_{01,i}. \quad (eq\ 5)$$

If the overall risks, or expected rates, of the solo effects (averaged over i) are R_{10} and R_{01} , respectively, then for the joint exposure it is:

$$R_{11} = R_{10} + R_{01} - E_i(R_{10,i}R_{01,i}), \quad (eq\ 6)$$

where, $E_i(R_{10,i}R_{01,i})$ represents the "expected" value (over i) of $R_{10,i}R_{01,i}$.

If the susceptibilities to the two exposures are mutually uncorrelated, then $E_i(R_{10,i}R_{01,i})$ equals $R_{10}R_{01}$. But since they may be correlated — to an unknown degree — and since the concern here is with rate criteria for independence, it is essential to understand the limits of $E_i(R_{10,i}R_{01,i})$. They can be shown to be

$$\begin{aligned} \max[E_i(R_{10,i}R_{01,i})] &= \min(R_{10}, R_{01}), \\ \min[E_i(R_{10,i}R_{01,i})] &= \max(0, 1 - R_{10} - R_{01}). \end{aligned} \quad (eq\ 7)$$

The application of these limits to formula 6 yields the boundaries (14):

$$\max(R_{10}, R_{01}) \leq R_{11} \leq \min(R_{10} + R_{01}, 1). \quad (eq\ 8)$$

Since the background risk ($R_{00,i}$) was assumed to be zero for each instance, the risks for the various exposure categories equal risk differences (RDs) relative to the domain of both exposures absent. Thus relation 8 may be recast as

$$\max(RD_{10}, RD_{01}) \leq RD_{11} \leq \min(RD_{10} + RD_{01}, 1). \quad (eq\ 9)$$

With the constraint in these formulations relaxed, suppose now that "spontaneous" transition (one in the absence of either exposure) occurs in proportion R_{00} of instances (ie, with rate R_{00}), and that R_{10}' , R_{01}' and R_{11}' are the frequencies of effect

in the domain of no spontaneous transition. Then the proportion of instances in which the joint exposure has effect, ie, the rate difference (RD) for it, is $RD_{11} = (1 - R_{00})R_{11}'$. For this, formula 8 implies the lower bound of $\max [(1 - R_{00})R_{10}', (1 - R_{00})R_{01}'] = \max (RD_{10}, RD_{01})$. The upper bound is $\min [(1 - R_{00}) (R_{10}' + R_{01}'), 1 - R_{00}] = \min (RD_{10} + RD_{01}, 1 - R_{00})$. This step leads back to the boundaries derived in the case of deterministic causation (expression 1) and to their corollaries for inference (expression 3). Similar arguments apply to the case of preventive factors (see example 4).

Implications

Definitions of causal and preventive interdependence were developed in this paper with a view to the essence of these phenomena only, without any regard to their implications in other areas; and since such interdependence can be inferred, with certain limitations, from epidemiologic data, there arises the question of what implications such inference might have.

The implications have to do with the origins of the interdependence (or independence) on one hand and with its consequences on the other, ie, the interest may be in inference about the interrelations of the two factors in the mechanisms of causation or prevention or in guidance for practical actions. My concern at this time is simply to illustrate each of these so as to clarify the *potential* for implications; any attempt at systematic analysis is beyond the scope of my present concerns (and abilities).

Inference about mechanisms

Various hypotheses about the relation of two factors in the causation (or prevention) of an event have express implications in terms of causal (or preventive) dependence. Thus the latter provides a basis for testing the former.

Consider, as a first example, two factors' "interchangeability in the same step of the causal process" (12), or of the preventive process. This may be interchangeability in the sense that each exposure is sufficient for accomplishing the component effect

("step") with no further effect from added dose, ie, interchangeability such as could be obtained between two types of penicillin regimen when each of them is appropriate insofar as the bacteria are penicillin-susceptible or between two platelet-active drug regimens with each of them fully effective in blocking adhesion. Such an interrelation means independence of effects, and the joint effect in the ultimate sense (cure of infection or prevention of coronary heart disease) is identical to each of the solo effects. Thus this kind of biologic interrelation means not only independence but an extremely high correlation of susceptibilities so that the rate manifestation corresponds to the respective boundaries of the ranges of independence (lower bound in expression 1 or upper bound in expression 3).

If the two exposures are interchangeable in the sense of having the very same function (as in the preceding case) but each represents a dose below that of maximal effect (in contrast to the preceding case), then the joint exposure can bring out instances of synergism, ie, instances in which the "combined dose" is needed to produce the effect. In this sense, two interchangeable exposures are synergistic as long as the response rate is monotonely increasing in the range at issue. By the same token, an exposure is always synergistic with itself in the range of such a dose-response pattern. The rate manifestation of the synergism, in the sense of the magnitude of the joint effect, may fall anywhere in the range from the larger of the two solo effects (positive or negative) to the maximal conceivable effect ($1 - R_{00}$ or $-R_{00}$ in causation and prevention, respectively).

The case in which the two factors "act in different steps in the process" (12) entails the need to distinguish between causal and preventive processes. For, if the two factors act in steps I and J, respectively, the process of causation involves completions of both steps I and J, whereas in prevention (see example 4) the blockage of either I or J will suffice. In the context of *causation* the effects are independent if the domains of incompleteness in I and J (in the absence of the factors at issue) are disjoint. This is an example of extreme negative correlation of

the susceptibilities, and its rate manifestation is given by the lower bound in expression 1. If the domains overlap, then the factors are synergistic. In the extreme of the latter, both solo effects are nil, while the joint exposure has an effect, so that the rate manifestation of synergism (expression 3) is extreme, rather than simply "multiplicative." If the two factors are *preventive* and the steps are of the all-or-none type, then their actions are independent, since in no instance is the blockage of both steps a prerequisite for effect; otherwise they can be synergistic.

The need to distinguish between causation and prevention applies also to the case in which the two factors act in parallel, redundant pathways. In causation the effects are independent, since "success" in both pathways is never needed. By contrast, in prevention the blockage of both pathways tends to be necessary in some instances, which means that the factors tend to be synergistic, an exception to this situation being the case of total disjointness of the domains of susceptibility to the solo effects.

If the two factors operate in mutually contradictory pathways, as is the case when the mechanisms involve, say, acidosis and alkalosis, respectively, then their actions are antagonistic.

Apart from the interrelations of the loci of the solo actions of the two factors, as already discussed, their causal (or preventive) dependence is determined by their direct effects on each other (when both are present), ie, on the *interaction* of the factors. Its consequence can be either enhancement or diminution of the component effects, and this occurrence, in turn, either accentuates or counteracts the dependence that would correspond to the absence of interaction. And when there would otherwise be independence, interaction of the factors can be the solo basis of synergism or antagonism.

Overall, it is apparent that the study of causal dependence (in epidemiology and elsewhere) can serve to detect patterns that are inconsistent with particular hypotheses about the interrelations of the sites of action. On the other hand, any given manifestation can arise as a consequence of several types of interrelation, and matters are complicated further by the

obfuscating roles of the relations of susceptibilities and the possible interaction between the factors.

Operational implications

Apart from scientific inferences about the nature of the causal (or preventive) processes, knowledge about causal (or preventive) interdependence has implications for particularistic judgements about the joint effect and, thereby, for operational decisions about seeking or avoiding the joint exposure.

The issues differ according to whether the solo effects are known for the instance at issue. When they are known, the case of greatest interest may be that in which both solo exposures are devoid of effect. In such an instance the joint exposure is without effect in the context of (simple) independence and (simple) antagonism, while in the case of synergism the possibility of effect remains. Thus, in the context of a desired effect, joint exposure holds promise despite solo failures, while with undesired effects the safety of solo exposures cannot be taken as assurance of the safety of joint exposure. In other types of instance as to known solo effects, the implications of causal (or preventive) interdependence are equally obvious.

When the solo effects are not known for the instance at hand, the characteristic question is whether the second factor has a greater or lesser effect in the presence of the first factor than when alone. Thus, in the context of health hazards, one tends to wonder whether the second exposure is more dangerous in the presence of the first than it is alone. With the effect, of an all-or-none kind, unknown in the instance at issue, the concern is with its probability of occurrence. But it is not totally obvious what probability criteria are appropriate for comparing the effects of the second factor between the presence and absence of the first one, because of the difference between the respective "baseline" probabilities or risks.

It does seem, however, that the equality of the effect of the second factor in the presence and absence of the first one can naturally be taken to mean that, of the instances in which the effect would remain absent without the second factor,

the latter would produce it in the same proportion in the presence and absence of the first factor. Formally, for causal factors, this criterion is

$$(R_{11} - R_{10}) / (1 - R_{10}) = (R_{01} - R_{00}) / (1 - R_{00}), \quad (\text{eq } 10)$$

where R_{10} is the risk of the event in the presence of the first factor alone, etc, as before. This criterion can be recast in terms of the corresponding risk differences as

$$RD_{11} = RD_{10} + RD_{01} - [(RD_{10})(RD_{01}) / (1 - R_{00})]. \quad (\text{eq } 11)$$

For preventive factors the equivalent of relation 10 is:

$$(R_{10} - R_{11}) / R_{10} = (R_{00} - R_{01}) / R_{00}, \quad (\text{eq } 12)$$

which implies:

$$RD_{11} = RD_{10} + (RD_{01})R_{10} / R_{00}. \quad (\text{eq } 13)$$

If RD_{11} "exceeds" (is more negative than) this value, then the effect of factor 2 is greater in the presence of factor 1 than in its absence, and, if RD_{11} falls short of this value, then the effect of factor 2 is less in the presence of factor 1 than in its absence.

These criteria for identity of effect are subject to epidemiologic evaluation without any reference to interdependence of effects; they are matters of comparative risks only.

On the other hand, these criteria correspond to independence on the condition of uncorrelated susceptibilities. By the same token, in the ordinary case of positively correlated susceptibilities, the presented criteria correspond to some degree of synergism.

Previous literature

In the previous epidemiologic literature the definition of causal or preventive interdependence ("interaction") or independence tends to be merged with its criterion in terms of epidemiologic data. Thus Rothman (9, 10, 11) has put forth a definition/criterion that is based on *average* risks corresponding to the absence of both factors, the solo presence of each, and the joint presence, respectively. Specifically, his definition and

criterion for causal independence has been the relation in equation 11 — or its approximation of $RD_{11} = RD_{10} + RD_{01}$ for the case in which the events is rare. By the same token, his definition for synergism is that RD_{11} exceeds the value implied by independence according to these formulas.

This amalgamation of ontological and epistemological concerns has been a source of confusion. One manifestation of this confusion is the recent tendency to distinguish between "public health" and other kinds of synergism or "interaction." Thus Blot & Day (1) think of synergism as defined by Rothman as a "public health concept," distinct from the statistical concept of interaction. The latter they seem to equate with the corresponding "biologic or mechanistic" concept, though without any clear rationale. Subsequently, Saracci (13) and Rothman et al (12) embraced the idea, except that they divided the biologic/mechanistic/statistical concept of Blot & Day — and others (5, 16) — into biologic and statistical "interaction." That Rothman's original definition/criterion for synergism has now become labeled "public health interaction" is justifiable in the sense that it deals with average risk, ie, with population parameters. But it is without justification as a concept or criterion of synergism because it does not express the ontological notion of cooperation of causes nor is it a derivative of it; instead, it is based on the statistical concept of independent events. (The latter concept Rothman applied on the population level, ie, in the sense of sampling results, in contrast to its appearance in the present paper — secondary to an ontological definition — in reference to the stochastic outcome in an individual instance.) None of these authors give an express definition for "biologic interaction," but some of them, at least, do recognize "statistical interaction" as merely a matter of style in the use of formal models for the occurrence of the event (the use of a nonadditive formulation).

From his definition/criterion for "synergism" Rothman has deduced the notion that synergism means coparticipation in a sufficient cause and vice versa (11). The latter proposition, too, is of the form of

a definition. Rephrased, it equates synergism with the existence of instances in which both factors are necessary for the event to come about, ie, it is the definition of "synergism" proposed in this communication. Yet Rothman's definition/criterion and the definition offered in the present paper are incongruous, the reason being that his deduction is a *non sequitur*. Rothman himself has dealt with a matter that should have brought out the discrepancy (9): a linear dose-response pattern (in the case of a rare event) means that component doses satisfy Rothman's criterion (of additivity of effects) for causal independence; on the other hand, the pattern implies the existence of instances in which both dose components are needed for the event to develop, ie, it implies the coparticipation of the components in a sufficient cause, ie, synergism in the sense of the putative corollary of the Rothman definition/criterion.

The Rothman "corollary," and thereby the present definition of causal interdependence, has been challenged by Koopman (4). He argues that "there are sources of interaction in addition to coparticipation in a sufficient cause." What he means by "interaction" remains obscure. If the meaning is nonadditivity in particular formulations of the model, neither Rothman nor I would take exception.

Rothman et al (12) have equated certain interrelations in the causal mechanism with additive and multiplicative risk functions, respectively, as has already been noted. The present analysis showed these propositions to be in error.

That decisions about joint exposure may be guided by rate differences for the solo and joint exposures has been well appreciated by Blot & Day (1) and, subsequently, by Saracci (13) and Rothman et al (12).

Overall, it may be noted that the essential difference between the previous epidemiologic literature and this communication is the point of departure. Until now the conceptualization of synergism has been burdened by epistemological and statistical concerns, whereas in this paper — and in a paper by Walker (15) — the take-off is guided by ontological considerations alone. With this formulation of the concept of causal and preventive interdependence, a good deal remains to be done

to clarify its utility in inferences about the relative points of action of a factor in the causal process and/or the interaction of the factor.

Acknowledgment

This research was supported by grant 5P01CA06373 from the National Cancer Institute.

References

1. Blot WJ, Day NE. Synergism and interaction: Are they equivalent? *Am j epidemiol* 110 (1979) 99—100. (Letter to the editor).
2. Greenhouse SW. Some epidemiologic issues for the 1980s. *Am j epidemiol* 112 (1980) 269—273.
3. Hewlett PS, Placbell RL. The interpretation of quantal response in biology. Edward Arnold, London 1979, chapter 7.
4. Koopman JS. Causal models and sources of interaction. *Am j epidemiol* 106 (1977) 439—444.
5. Kupper LL, Hogan MD. Interaction in epidemiologic studies. *Am j epidemiol* 108 (1978) 447—453.
6. Miettinen OS. Confounding and effect-modification. *Am j epidemiol* 100 (1974) 350—353.
7. Miettinen OS. Estimability and estimation in case-referent studies. *Am j epidemiol* 103 (1976) 226—235.
8. Miettinen OS, Cook EF. Confounding: Essence and detection. *Am j epidemiol* 114 (1981) 593—603.
9. Rothman KJ. Synergy and antagonism in cause-effect relationships. *Am j epidemiol* 99 (1974) 385—388.
10. Rothman KJ. The estimation of synergy and antagonism. *Am j epidemiol* 103 (1976) 506—511.
11. Rothman KJ. Causes. *Am j epidemiol* 104 (1976) 587—592.
12. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am j epidemiol* 112 (1980) 467—470.
13. Saracci R. Interaction and synergism. *Am j epidemiol* 112 (1980) 465—466.
14. Wahrendorf J, Brown CC. Bootstrapping a basic inequality in the analysis of joint action of two drugs. *Biometrics* 36 (1980) 653—657.
15. Walker AM. Proportion of disease attributable to the combined effect of two factors. *Int j epidemiol* 10 (1981) 81—85.
16. Walter SD, Halford TR. Additive, multiplicative and other models for disease risks. *Am j epidemiol* 108 (1978) 341—346.
17. ———. *Dorland's illustrated medical dictionary*. 23rd edition. WB Saunders Co, Philadelphia & London 1957.
18. ———. *Webster's new collegiate dictionary*. G & C Merriam Co, Springfield, MA 1977.

Received for publication: 28 January 1982