

Appendix II

Modelling complex interactions in longitudinal population models

Introduction¹

Generic population health models [1] offer the possibility of easy-to-use longitudinal models that can be used for population health measurement. In many cases, generic models can be implemented with generally available epidemiologic and demographic data and can be used to quantify real population health risks. In spite of these possibilities, analysts who are faced with the need to model complex disease processes, such as those presented by many chronic diseases, for example, may find that four health states are not sufficient for a decent level of realism. Typically chronic disease processes can involve an arbitrary number of disease stages accessed in a progressive, non-reversible fashion. Or, if an analyst is interested in modelling multiple risk factors and disease states, the factor limiting the realism of the model may not be the number of states available in the model but rather the number of interactions between risk factors, in other words, the need to model combinations of risk factors requiring the definition of many more than four states. For such situations, four different health states is far too few and the use of generic models [1] will therefore ultimately be unsatisfying for some population health problems and a greater level of heterogeneity will be sought. Although different approaches are possible, and many of them will share a similar set methods for addressing the problem we describe, microsimulation is one natural approach. With microsimulation methods, one can readily investigate questions of multiple interacting risk factors and causes of disease (i.e. “causal web”). However, the technology of microsimulation does not change the fundamental need to characterize a joint distribution from a set of observations of marginal distributions. Moreover, the question of characterizing in quantitative terms the interactions in the causal web becomes important. Thus, the kinds of data transformation that are required to define the initial state of the population simulation can be more complex. Several generic insights are nevertheless possible.

When one wishes to model several population states in a longitudinal

1 I am grateful to Claude Charette and Steve Gribble of Statistics Canada for their assistance in helping me to start thinking about the problem discussed in this appendix and for their having worked through a complete example.

population model, lack of information about the joint distribution of risk (either longitudinal, in terms of rates, or cross sectional, in terms of prevalences) is a common problem. Even for the simple case of estimating the effect of a risk factor within a given strata of the population, in the absence of knowledge about the joint distribution of risk factors within the stratum, the estimation of effects is – in most cases – hopelessly confounded [2]. Of course, if risk distributions were statistically independent, the problem would be trivial since the joint distribution would merely be calculated as the product of the marginal distributions. Yet departures from independence are often precisely what make the joint distribution of risk of interest [2]. So determining the joint distributions of non-independent risk factors is not only an ubiquitous but also an important problem.

These questions can be related more generally to concepts of "interaction" as discussed in the epidemiological literature [3-22]. However, in the mainstream literature, interaction is typically defined with respect to rates, relative rates or incidence proportions [3]. In any case, since rates, relative rates and relative incidence proportions are measured longitudinally, one can conclude that the kind of interaction typically discussed in the epidemiological literature is what could be called "longitudinal interaction" (i.e. related to interaction in rates). Despite its similar importance, the problem of determining the joint distribution of risk factors and stratification variables in cross section typically goes underrecognized, although it has been mentioned as a source of unavoidable estimation bias [19,21].

In either the longitudinal or the cross-sectional case, however, it is important to be modest in trying to relate statistical concepts of dependence and independence to biological processes of interaction [6]. It is important to keep in mind that, among other problems in moving from a statistical to a biological concept, that a variety of different biological interactions can produce the same statistically observed outcomes[13]. Recognizing the inherent gulf between biological and statistical concepts, we develop here a pragmatic approach to "interaction" that is defined primarily in statistical terms. Our concept will therefore be mostly agnostic about the true underlying biological mechanisms, although it will occasionally be possible to rule out certain biological mechanisms as inconsistent with the observed data.

Moreover, in the approach presented here, headway can be made on both the problem of longitudinal and cross-sectional interaction through the application of a single conceptual framework and approach. Indeed, since the current cross-sectional distribution of exposure is merely the cumulative result of past events, it should in principle be possible to approach the problem of the joint prevalence of risk factors (i.e. static, or cross-sectional, interaction) as well as the decomposition of aggregate rates (i.e. dynamic, or longitudinal, interaction) with a common set of methods.

For both conceptual clarity and measurement purposes, however, it is useful to distinguish dynamic and static concepts of interaction in the following exposition, where we define a minimal set of necessary and sufficient information to determine the joint distribution in the static case as well as solve the decomposition of aggregate rates. We specify that the "joint distribution" can be understood to refer either to the joint distribution of persons or of person time and that the methods employed do not depend on one or the other distributional concept, although it is of course always useful to use a consistent set of definitions in the course of a particular problem. It is also useful to recognize, at least as an approximation, that the concepts of prevalence, rates and odds can be referred to together, as necessary, by means of the common term "risk" [3] without damage to either common sense or the main conclusions of the argument.

Static problem

To fix ideas, we refer to a case of stratification variables and risk factors, although the actual boundaries between the two concepts are fluid. Typical stratification variables might be age, sex, income or educational status, whereas typical risk factors are usually the exposures of primary interest in the question at hand. However, these designations are mainly arbitrary since risk factors will typically become stratification variables once they are well known. In any case, suppose there are K strata for the stratification variable $\underline{M} = \{M_k\}$, $k = 1, \dots, K$, and J levels of risk factor $\underline{X} = \{X_j\}$, $j = 1, \dots, J$. It should be clear that we refer here to the partition of the population into a 2-way discrete classification.

We may start by noticing that, in such a case, it will always be possible to define the joint distribution of risk within strata as the solution of a system of $K \times J$ independent equations. For the problem to be well posed, these equations, whatever their actual positive content, must be written in terms of a vector of the unknown joint prevalences $\mathbf{p}$, a matrix of known coefficients $\mathbf{A}$ and a vector of observable constants $\mathbf{b}$. The observable constants $\mathbf{b}$ will typically be either the marginal prevalences or zeros (discussed below), and the coefficients $\mathbf{A}$ will typically be either known relative risks or other known constants (discussed below). We return to this in more detail subsequently. In any case, based on a suitable identification of the different elements of the system, we then write $\mathbf{A}\mathbf{p} = \mathbf{b}$. It is well known that whenever the matrix $\mathbf{A}$ is of full rank, the system has a unique solution for $\mathbf{p}$.

To define solution conditions that are consistent with the empirical problem, it is reasonable to assume that the marginal distributions $\mathbf{b}$ of both X and M are known. The marginal distributions can be thought of here as the population-level risk of (i.e. the probability of finding) the members of the different risk-factor and stratification-variable levels based on uniform random sampling from the entire population. Actual observations of $\mathbf{b}$ will typically not strictly fulfill these conditions but rather approximate them.

These marginal prevalences, which we assume are observed, define $J + K - 1$ independent summing-up equations. That is, for each marginal prevalence, we can always write K equations of the form $m_i = \sum_{j=1}^J p_{i,j}$, as well as for each risk factor J equations $x_j = \sum_{i=1}^K p_{i,j}$. Note that $p_{i,j}$ are unknown in both cases. When written in matrix form, these $J + K - 1$ equations define a submatrix of $\mathbf{A}$ whose coefficients are all ones (1s). Note however that although we could in principle write $J + K$ equations, only $J + K - 1$ of the equations are actually independent, since one of the equations can always be expressed in terms of a linear combination of the others (i.e. by summing across either risk-factor or stratification-variable levels, the total population is recovered: $1 = \sum_{j=1}^J p_{i,j} = \sum_{i=1}^K p_{i,j}$).

Since the observed marginal prevalences yield only $K + J - 1$ independent equations in terms of the joint prevalences, $(K \times J) - (K + J - 1) = (J - 1) \times (K - 1)$ additional equations are required in order for the system to be solvable. This number $(J - 1) \times (K - 1)$ is evidently equal to zero only when K or $J = 1$.

We propose that, in general, it will be reasonable to characterize these $(J-1) \times (K-1)$ additional equations in terms of relative risk relations. The advantage of relative risk measures for this purpose are several. First, and most importantly, relative risk provides a concise way to define interaction in statistically observable terms. Second, estimates of relative risk as defined below are typically available from case-control studies. Third, it will often be possible to make reasonable guesses about the magnitude of a given relative risk even in the absence of data. Finally, using relative risk measures such as those defined below to construct the required additional equations to complete the system $\mathbf{A}\mathbf{p} = \mathbf{b}$ does not for the moment require any assumption about the functional form of relative risk (e.g. additive, multiplicative etc.). We return to this topic when we discuss dynamic interaction, below.

As is typical, we define relative risk as a ratio of conditional risk

$$\text{Eq. (1)} \quad \wp_{k,1} \equiv \frac{p_k(X = x_j)}{p_1(X = x_j)},$$

where p_k represents the probability of finding an exposure unit (person or person-time) in stratum level k conditional upon a specific exposure status (e.g. $X = x_j$). We say that when there is interaction between levels of the risk-factor and the stratification-variables $\wp_k \neq 1$. Subsequently, we define more precisely what kind of conditional risk measures are required to solve $\mathbf{A}\mathbf{p} = \mathbf{b}$.

Suppose, for example, that, for each risk-factor level j , an equation of the form

$$\text{Eq. (2)} \quad \wp_k = \frac{p_k(X = x_j)}{p_1(X = x_j)}$$

can be written, where $X = x_j$ signifies that there is exposure to risk factor level j . This represents the typical sort of measure that is available from case-control studies. Suppose for the moment that there are two stratification-variable levels ($k = 1, 2$), one risk factor level, one baseline, or non-exposed, level, and that the interaction of risk-factor level j with stratum 2 is written $\wp_{2,j}$

$$\text{Eq. (3)} \quad \wp_{2,j} = \frac{p_{2,j}/m_2}{p_{1,j}/m_1},$$

where the marginal prevalence of measurement units in stratum 1 is given by m_1 , the marginal prevalence of measurement units in stratum 2 is given by m_2 and we write $p_{2,j}$ and $p_{1,j}$ for the unknown joint prevalence of risk-factor level j in stratum-level 1 and 2, respectively. Rearranging the latter equation for the relative risk, we can then evidently write

$$\text{Eq. (4)} \quad p_{1,j}m_2\rho_{2,j} - p_{2,j}m_1 = 0,$$

which is a linear equation in terms of the unknown joint prevalences $p_{2,j}$ and $p_{1,j}$. On the basis of the observation of the value $\rho_{2,j}$, such an equation can be added to the $(K + J - 1)$ equations that were previously written in terms of the summing-up constraints for the marginal prevalences.

Since, when the marginal prevalences are known (which is our assumption), it is always possible to write the summing-up constraints, and since each relative risk $\rho_{i,j}$ that we know defines one additional equation in terms of a subset of the joint prevalences $p_{i,j}$ and some of the marginal prevalences m_i and x_j , it will always be possible to solve for the full set of joint prevalences whenever we know the value of a set of $(J-1) \times (K-1)$ complete and independent relative risks. The equations are shown in matrix form for the $J = K = 3$ case in the Figure. The solution can be obtained by matrix algebra or by Gaussian elimination (i.e. repeat substitution).

Figure: The matrix equations for the $J = K = 3$ case.

$$\begin{pmatrix} m_1 & -m_3\rho_{3,b} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -m_2\rho_{2,b} & m_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & m_1 & -m_3\rho_{3,c} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -m_2\rho_{2,c} & m_1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} p_{3,b} \\ p_{1,b} \\ p_{2,b} \\ p_{3,c} \\ p_{1,c} \\ p_{2,c} \\ p_{1,a} \\ p_{2,a} \\ p_{3,a} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x_b \\ m_1 \\ m_2 \\ m_3 \\ x_c \end{pmatrix}$$

What is an “independent” or “complete” relative risk in this context? Each relative risk relates the conditional prevalence of a single stratum-risk pair with that of another stratum-risk pair. Independence simply requires that there are no redundant stratum-risk pairs among the available set of observed relative risks. Completeness, which is in fact satisfied whenever there are $(J-1) \times (K-1)$ independent relative risks, requires that all existing strata and risk levels can be mutually related, for example, for four strata, three stratum-risk pairs are required to mutually relate all four strata. For example, if the strata are labeled i, j, k , and l , any of the sets of strata pairs $\{i \rightarrow j, j \rightarrow k, k \rightarrow l\}$, $\{i \rightarrow j, j \rightarrow k, j \rightarrow l\}$ and $\{i \rightarrow j, j \rightarrow k, i \rightarrow l\}$ are complete. A complete set of strata pairs is required for each risk level except for “baseline” risk since the relative risk of the “last” risk factor level can always be expressed as in terms of the other relative risks.

So far we have discussed the case of a single risk factor with J levels $\underline{X} = \{X_j\}$, $j = 1, \dots, J$ and a single stratification variable with K levels $\underline{M} = \{M_k\}$, $i = 1, \dots, K$. However, when there are additional risk factors or stratification variables that are also of interest, provided there are enough additional relative risks estimates available, the approach described above can be iteratively applied to find the solution to the joint prevalence. To begin, we notice first that additional risk factors (or stratification variables) must be either interacting or non-interacting with the initial pair of risk factor and stratification variable. If the additional variables are non-interacting, a new joint distribution can easily be obtained by multiplying each element of the joint distribution solved for in the first step by each element of the marginal distribution of the new risk factor (or stratification variable).

However, if the new risk factor or stratification variable is interacting, it must interact either with the existing risk factor, with the existing stratification variable or with both. If it interacts with only one of them, the process of writing summing-up constraints and relative-risk equations can be repeated with that one. If, on the other hand, it interacts with both of them, then one of the interactions can be chosen first. The new variable must interact either with levels, or with aggregates of levels, of the marginal distribution of the chosen variable. On the basis of that interaction (i.e. using the relative risk observations that are available), one then write a

new system of the form $\mathbf{A}\mathbf{p} = \mathbf{b}$, which will again be solvable whenever the matrix $\mathbf{A}$ can be written so as to be of full rank. The same process can then be repeated for the other interaction.

Provided the conditions for the existence and uniqueness of a solution are satisfied for each step (i.e. the matrix $\mathbf{A}$ is of full rank), this process can be repeated an arbitrary number of times. This is true even when none of the joint prevalences that are sought are actually known to start with. Moreover, because of considerations of symmetry, regardless of which risk-factor/stratification-variable pair is chosen as the starting point (assuming there are several alternatives), the end-point of the iterated solution process must be the same (i.e. in practice, the solution is constrained by the available data on marginal distributions and relative risk).

One may further observe that, due to measurement error or bias in measurement, the supposedly known quantities can in fact be uncertain. As a result, the mapping from the space of the known quantities (marginal prevalences) to the space of the joint prevalences may be subject to uncertainty. In this case, a notional uncertainty analysis can proceed as follows. Suppose error in the known quantities is normally distributed. Since the solution vector of joint prevalence is, according to matrix theory, a linear combination of the known quantities, its elements will also tend to have normal distributions. This tendency will be more pronounced the more independent known quantities are represented in the solution expressions.

Dynamic problem

As noted earlier, the epidemiological literature has typically been concerned what we have termed dynamic interaction (interaction in rates).

Interaction hypotheses

In the static case, we have discussed interaction without regard to its functional form (cf. Eq. 3, which depends on the magnitude of interaction but not on its specification). This sort of measure is non-parametric in the sense that whatever kind of interaction is present, it is captured in the data that are observed. This type of assumption is consistent moreover with the sort of observations of static interaction that are typically available

from case-control studies. In the case of dynamic interaction, however, considerable attention has been devoted to a number of different canonical forms of interaction and their implications [3,11,14-16].

Additive models

For example, a simple additive model for rates has been defined as follows [3]

$$\text{Eq. (5)} \quad r_k(\underline{x}) = \alpha_k + \underline{x}\beta,$$

where $\underline{x}$ is the vector of exposures, k indexes the stratum and β is a vector of effect measures expressing the change in risk per unit increase in the exposure $\underline{x}$. Since $r_k(\underline{0}) \equiv \alpha_k$, the additive model implies (or can be derived from) additivity of the rate difference

$$\text{Eq. (6)} \quad D_k(x) = r_k(\underline{x}) - r_k(\underline{0}) = \underline{x}\beta.$$

Since the model of additive rate-difference is synonymous with the hypothesis of “no causal interaction” under Rothman’s sufficient-component cause model [3], the lack of causal interaction can be taken as a definition of the simple model of additive rate difference, and a finding of additive rate difference can be taken as evidence in favour of the absence of causal interaction.

Another model typically referred to as “additive” is the additive relative rate model. The rate model that corresponds to the additive relative-rate model can be written as follows [14,16]

$$\text{Eq. (7)} \quad r_k(\underline{x}) = \alpha_k(1 + \underline{x}\beta^*),$$

where β^* is a vector of constant effect measures expressing the change in relative risk per unit increase in $\underline{x}$ (note that this is a different measure of effect than that which is relevant to the model of additive rate difference, above). The additive relative rate model implies that excess relative risk, defined as

$$\text{Eq. (8)} \quad E_k(\underline{x}) \equiv \frac{r_k(\underline{x})}{r_k(\underline{0})} - 1 = \underline{x}\beta^*$$

is also additive. The additive relative rate model is also a proportional risk model [3], since, as can be seen from the definition of the model, stratum-specific exposure effects differ only by a constant of proportionality α_k . The

additive relative rate model is therefore is not a model of non-interaction (as has sometimes been claimed) but rather a restriction on interaction to a particular multiplicative form (i.e. constant of proportionality α_k).

A less restricted additive model is the following [3]

$$\text{Eq. (9)} \quad r_k(\underline{x}) = \alpha_k + x\beta_k.$$

Here, the effect of exposure is stratum specific, and interaction is not limited to a particular mathematical specification. Such stratum-specific interaction would therefore need to be observed. This less restricted additive model can also be written in terms of an additive relative (or excess) risk model

$$\text{Eq. (10)} \quad r_k(\underline{x}) = \alpha_k + x\beta_k = \alpha_k \left(1 + \frac{x\beta_k}{\alpha_k}\right) \equiv \alpha_k (1 + x\hat{\beta}_k).$$

Here, each element β_k is divided by the corresponding α_k and renamed $\hat{\beta}_k$, yielding an additive relative risk model with stratum specific coefficients. This is in fact the specification in which additivity of risk is usually tested for in studies [3].

In every additive model but the simple additive model risk-difference model of no interaction, therefore, the assumed additivity relation in underlying rates, risk or odds can be represented in terms of a corresponding additivity relation in relative risk, relative rates or relative odds. Moreover, the different additive models represent a spectrum of (additive) specifications for interaction: the least restricted model of Eq. 9 allows stratification variables to interact with risk factors in a non-specified manner; the model of additive relative risk allows risk factors to interact with stratification variables but limits the interaction to constant factors α_k ; finally, the additive risk model implies no interaction between stratification variables and exposures [3].

Equivalence between additive models

The simple additive model for rates (Eq. 5) is equivalent to an additive model for relative rates (Eq. 7) only if [3]:

- there is only one stratum, or
- the baseline rate $r_k(\underline{0})$ is constant across strata (i.e. $r_k(\underline{0}) = r(\underline{0})$).

However, we have shown that it is always possible to solve for unique values for $r_k(\underline{x})$ when $J + K$ marginal prevalences are known and there are $(J-1) \times (K-1)$ independent known relative risks. We now show that, in general, the existence of such a solution will almost always require baseline risk to differ across strata, implying that the additive model for rates is almost never equivalent to the additive model for relative rates.

The stratum-specific (i.e. joint) risk, rate or odds that are the solution of the matrix equation correspond to a basis vector in the space of the joint prevalences. Moreover, precisely because the solution vector for the joint prevalences is a basis in the space of joint prevalence, each solution element is necessarily algebraically different (since if they were the same, one element would be redundant and, contrary to the construction of the problem, the basis would then not span the space of joint prevalence). We can nevertheless ask under what conditions these algebraically different solution elements may be numerically equal.

The conditions for this can be determined by setting the different solution expressions for baseline risk equal, $r_{i,0} = r_{j,0}$ for all possible i,j pairs, where i and j index the stratum levels. For K levels, the number of possible i,j pairs where $i \neq j$ can be written compactly as the binomial coefficient $\binom{K}{2}$. When the relative risks are independent and complete, solutions for baseline risk will moreover necessarily be expressed in terms of all $J \times K$ known quantities for marginal risks and relative risks. The baseline risks cannot be expressed in terms of fewer than $J \times K$ known quantities because solvability of the system requires each of the known quantities and no fewer. Each of the $\binom{K}{2}$ equalities that need to be written to enforce equality of baseline risk across strata is therefore an expression written in terms of each of the $J \times K$ known quantities. Since each equality can be solved in terms of one of the $J \times K$ known quantities, each of the $\binom{K}{2}$ equality constraints induces a corresponding equality constraint on each of the $J \times K$ other known quantities it contains.

Since there are $\binom{K}{2}$ equalities that can be solved for a given known quantity, these in turn induce $\binom{\binom{K}{2}}{2}$ pairwise equalities in terms of the remaining $J \times K - 1$ known quantities. From this group of pairwise constraints, another

one of the $J \times K - 2$ remaining known quantities can then be solved for. However, this solved-for quantity now has to satisfy $\left(\left(\begin{smallmatrix} K \\ 2 \end{smallmatrix}\right)\right)_2$ constraints expressed in terms of the remaining $J \times K - 3$ known quantities. Proceeding in this fashion, it can be seen that the binomial coefficient $\left(\begin{smallmatrix} K \\ 2 \end{smallmatrix}\right)$ must be nested $J \times K$ times, in other words, until all the known quantities except for the last one have been expressed in terms of all the others. For the case of 4 strata and 4 known quantities, the first known quantity therefore has to satisfy 6 constraints, the second 15, the third 105 and the last 5460. However, since the order of choosing the known quantities is arbitrary, symmetry requires that each of the known quantities should satisfy 5460 additional constraints.

However, since each such additional constraint represents a linear combination of the solution values for the joint risk, rate or odds, even one such additional constraint is too many for the known quantities to satisfy. This is because the only combinations of the solution values for joint risk, rate or odds that also are combinations of the known quantities are those specified in the system of equations $\mathbf{A} \cdot \mathbf{p} = \mathbf{b}$, which was solved to obtain the joint prevalences. In fact, the matrix $\mathbf{A}$ from that system of equations is guaranteed by construction (i.e. by the full rank condition) to represent a unique transformation that maps elements from the space of the joint prevalences to the space of the known quantities.

If any single additional linear combination was in fact possible, it would be equivalent to adding an extra row to the original matrix equation. However, since the original matrix equation is the unique transformation that maps the joint prevalences into the known quantities, two possibilities arise: either the additional row is linearly dependent on the other, original, rows, in which case the matrix is singular (and the “additional” constraint is therefore implicit in the original system and is consequently not an additional constraint), or the additional row is independent (and the matrix is therefore inconsistent). In either case, the full rank condition is violated and a solution to the matrix equation is impossible. Therefore, any additional constraints on the known quantities in terms of the joint solved-for prevalences are not possible.

There is nevertheless one case where the condition of equality of baseline risk does not impose any additional constraints on the known quantities. That is the case of two strata only (regardless of the number of risk levels). Since $\binom{2}{2} = 1$, equality of baseline risk across the two strata can induce only a single equality constraint on the joint solution quantities, but, since $\binom{2}{2} = 0$, no equality constraints are in turn induced on the marginal prevalences).

It has already been noted in [3] that the additive relative risk model is equivalent to the additive risk model only when there is only one stratum or when baseline risk is equal across strata. The arguments presented here provide an important restriction on the second possibility, namely, that it is only possible in principle for baseline risk to be equal across strata when there are only two strata.

The additive relative risk model can therefore only be equivalent to the simple additive risk model of no interaction when there are at most two strata (and it is by no means guaranteed to hold in the two-strata case). In any other case, the additive risk model cannot be equivalent to the additive relative risk model.

Multiplicative model

A multiplicative model for risk is typically written [21]:

$$\text{Eq. (11)} \quad r_k(\underline{x}) = \exp \beta(\underline{x}) \cdot r_k(\underline{0}),$$

from which it can be seen that relative risk can be written:

$$\text{Eq. (12)} \quad R_k(\underline{x}) \equiv \frac{r_k(\underline{x})}{r_k(\underline{0})} = \frac{\exp \beta(\underline{x}) \cdot r_k(\underline{0})}{\exp \beta(\underline{0}) \cdot r_k(\underline{0})} = \frac{\exp \beta(\underline{x})}{\exp \beta(\underline{0})}.$$

Since by Eq. 11, $\exp \beta(\underline{0})$ must be equal to 1, it follows that

$$\text{Eq. (13)} \quad R_k(\underline{x}) = \exp \beta(\underline{x}),$$

which is a multiplicative model for relative risk. The multiplicative relative risk model is of considerable interest in practical work since estimates of relative risk obtained using such a model are interpretable as standardized parameters, i.e. parameters standardized to the risk distribution in the study population from which they were obtained [21].

Multiplicative versus additive models of interaction

While it has been claimed that additivity of risk, rate and odds remains the “most basic manifestation of independent action” [6], empirical additivity of rates does not by itself imply the absence of interaction [3] and it is therefore a necessary but not sufficient condition for a definite claim about the absence of interaction. Interaction, moreover, as we have discussed above, can have different specifications, including additive and multiplicative relative rate specifications.

In any case, for certain simple biological models of interaction, it is in general possible to say whether an additive or multiplicative model is preferred. For example, in simple multi-stage models model derived from studies of carcinogenesis, it can be shown that two risk factors acting on the same stage of pathogenesis interact in an additive fashion, whereas risk factors that act on different stages interact multiplicatively [10, 12].

For example, a simple carcinogenesis model posits that a single cell, by undergoing a series of transformations, results in a cell capable of malignant replication; such a model is consistent with the well known power-law relationship for cancer incidence [10]. Since noxious exposures increase the probability of transformations, when they facilitate separate stages of the transformation to malignancy their joint effect combines multiplicatively, but when they facilitate the same stage, the joint effect combines additively. A multi-stage multiplicative model has solid empirical support for non-hormonally dependent cancers of epithelial origin [10].

Therefore, when pathogenesis is multi-factorial and exposures affect different stages of disease, a multiplicative model appears to be the canonical model, whereas when exposures affect the same stage of disease occurrence, an additive model may be canonical. Whether the additive model is an additive rate or risk model will depend on the formulation of the underlying disease model, i.e. in terms of either rates or relative rates and, as we have shown, an additive rate model will almost never be an additive relative rate model. In any case, the additive relative rate model is of some interest because of its use as a common model to test for additivity.

Finally, additive and multiplicative relative rate models cannot be equal: each specification has different quantitative implications for outcomes. In particular, a multiplicative model of interaction will in general result in a higher rate of disease than any additive model.

Since “it is difficult to imagine situations in which typical stratifying factors ... would be unassociated with any causal cofactors” of the exposures $\underline{x}$, we propose that the additive relative rate model and the multiplicative relative rate model may be representing two different possible forms of interaction that respond to different possible models of biological interaction in multi-stage disease causation. As we shall see in the following section, this characterization offers the additional advantage that both cases can be characterized in terms of relative risk only.

Rate decomposition

Now we consider transition events, in particular from disease-free to diseased status. Transition events are modelled by rates, and are therefore fundamentally longitudinal. Here, to fix ideas, we focus on the case of hazards, i.e. instantaneous rates.

The total population rate of transition to disease y is equal to the per-capita rate of transition, r^y , multiplied by the prevalence of the group without disease y , p^0 (i.e. the population prevalence of the group capable of experiencing the transition), in other words, the quantity $r^y \cdot p^0$. We write p with superscript 0 to emphasize the prevalence of the group without disease y , and we write r with superscript y to emphasize the rate of transition to y , respectively. Suppose that the disease-free population is furthermore divided into subgroups by risk-factor and stratification-variable status according to the methods of the previous section regarding the static problem. The total population transition rate $r^y \cdot p^0$ must then necessarily be equal to the sum of the products of per-capita rates in the different subgroups and the prevalences of the respective subgroups. Suppose that the subgroups are denominated by subscripts denoting stratum/risk-factor levels and that we write superscripts to emphasize status with respect to disease y . When there are two strata (1, 2) and two risk-factor levels (a , b), we can then write

$$\text{Eq. (14)} \quad r^y \cdot p^0 = r_{1,a}^y \cdot p_{1,a}^0 + r_{2,a}^y \cdot p_{2,a}^0 + r_{1,b}^y \cdot p_{1,b}^0 + r_{2,b}^y \cdot p_{2,b}^0.$$

By the methods of the static problem, above, the prevalence values in an equation of this type are assumed to be known. Since it is the sort of quantity which is in principle observable, we further suppose here that the total population rate, $r^y \cdot p^0$, is also known. However, since there are $K \times J$ unknown subgroup rates, for the moment, we have only one equation with $K \times J$ unknowns. In the absence of $(K \times J) - 1$ additional independent equations, it is clearly impossible to solve for the unknown subgroup rates. This is a substantially worse situation than that which was faced in the static problem, where we required only $(K - 1) \times (J - 1)$ additional equations, which we supposed were available in the form of relative risks. In other words, here we need $K + J - 2$ more equations than we needed to solve the static case. What can we do to make the information demands in the dynamic case in practice no greater than they were in the static case?

One way is to suppose that we know only a relatively small number of relative rates but that we know that relative rates follow a particular mathematical form, such as additive or multiplicative. In such a case, we can then solve for the different subgroup rates. In other words, if we know only $(K - 1) + (J - 1)$ independent relative rates *and* we know a rule $f(\cdot)$ for combining relative rates, we can easily derive the additional relative rates required to make the system solvable. To continue the dynamic example we have started, above, if we know the following two relative rates

$$\text{Eq. (15)} \quad \wp_{1 \rightarrow 2, a} = \frac{r_{2, a}}{r_{1, a}}, \wp_{1, a \rightarrow b} = \frac{r_{1, b}}{r_{1, a}},$$

then we can easily derive the third required relative rate

$$\text{Eq. (16)} \quad \wp_{1 \rightarrow 2, a \rightarrow b} = \frac{r_{2, b}}{r_{1, a}} = f(\wp_{1 \rightarrow 2, a}, \wp_{1, a \rightarrow b}).$$

If relative rates are multiplicative, the derived relative rate is calculated

$$\text{Eq. (17)} \quad \wp_{1 \rightarrow 2, a \rightarrow b} = \wp_{1 \rightarrow 2, a} \cdot \wp_{1, a \rightarrow b},$$

but if relative rates are additive, the derived relative rate is calculated

$$\text{Eq. (18)} \quad \wp_{1 \rightarrow 2, a \rightarrow b} = (\wp_{1 \rightarrow 2, a} - 1) + (\wp_{1, a \rightarrow b} - 1) + 1.$$

Following the approach for the static case, with these three relative rates it is possible to write enough equations in the four unknown subgroup-specific rates to solve for all the subgroup rates.

While in general it is impossible to say with any confidence what is the correct relative risk law to apply, broad statements of the following type appear to be not unreasonable:

- The additive rate model is unlikely to apply since there will in general be interaction (i.e. the statistical association of stratification and exposure variables).
- Interaction can be characterized according to two canonical forms in terms of relative rates (additive and multiplicative).
- Each form has specific implications for outcomes, and each can be related with profit to a specific part of the epidemiological literature.
- The multiplicative form for interaction of relative risk implies a larger subgroup risk than the additive form of interaction.
- The multiplicative form of interaction has the interpretation that exposures operate at different stages of pathogenesis, or, in other words, through different causal pathways.

References

1. Lauer JA, Roehrich K, Wirth H, Charette C, Gribble S, Murray CJL. PopMod: a longitudinal population model with two interacting disease states. *Cost effectiveness and resource allocation*, 2003, 1:6.
2. Greenland S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *Int J Epidemiol* 2001;**30**:1343-50.
3. Greenland S. Additive risk versus additive relative risk models. *Epidemiology*. 1993; 4:32-36.
4. Walker AM Rothman KJ (1982) Models of varying parametric form in case-referent studies. *Am J Epidemiol* 115, 129-137.
5. Greenland S. Basic problems in interaction assessment. *Environmental health perspectives supplements*, 101, supplement 4, 59-66.
6. Greenland S, Poole C. Invariants and non-invariants in the concept of interdependent effects. *Scan J Work Environ Health*, 1988;14:125-129.
7. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am J Epidemiol*, 1980; 112:467-470.
8. Miettinen OS. Causal and preventive interdependence: elementary principles. *Scan J Work Environ Health*, 1982;8:159-168.
9. Weinberg CR. Applicability of the simple independent action model to epidemiologic studies involving two factors and a dichotomous outcome. *Am J Epidemiol*, 1986; 123: 162-173.
10. Breslow NE, Day NE. *Statistical methods in cancer research, vol. I. The analysis of case-control studies*. Lyon, International Agency for Research on Cancer, 1980.
11. Breslow NE, Day NE. *Statistical methods in cancer research, vol. II. The design and analysis of cohort studies*. Lyon, International Agency for Research on Cancer, 1987.

12. Doll R (1971) The age distribution of cancer: implications for models of carcinogenesis. *J R Stat Soc Lond series A*, 134, 133-156.
13. Siemiatycki J Thomas DC (1981) Biological models and statistical interactions: an example from multi-stage carcinogenesis. *Int J Epidemiol*, 10, 383-387.
14. Moolgavkar SH, Venzon DJ. General relative risk regression models for epidemiologic studies. *Am J Epidemiol*, 1987; 126: 949-961.
15. Thomas DC. General relative-risk models for survival time and matched case-control analysis. *Biometrics*, 1981; 37:673-686.
16. Breslow NE, Storer BE. General relative-risk functions for case-control studies. *Am J Epidemiol*, 1985; 122:149-162.
17. Breslow NE (1976) Regression analysis of the log odds ratio: a method for retrospective studies. *Biometrics*, 32, 409-416.
18. Holford TR (1980) The analysis of rates and of survivorship using log-linear models. *Biometrics*, 36, 229-306.
19. Maldonado G, Greenland S. A comparison of the performance of model-based confidence intervals when the correct model form is unknown: coverage of asymptotic means. *Epidemiology*, 1994; 5: 171-182.
20. Maldonado G, Greenland S. Interpreting model coefficients when the true model form is unknown. *Epidemiology*, 1993; 4: 310-318.
21. Greenland S, Maldonado G. The interpretation of multiplicative model parameters as standardized parameters. *Statistics in Medicine*, 1994; 13: 989-999.
22. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. *Am J Epidemiol*, 1993; 138: 923-936.

