Epidemiology Review

Effect Modification in Epidemiology and Medicine

Farin Kamangar MD PhD•[1,2]

Abstract

Effect modification, also known as interaction or heterogeneity of effect, is an important concept in epidemiology. This article reviews the definition and types of effect modification, methods to detect effect modification, the reasons for observing effect modification in epidemiologic studies, the importance of choice of model in finding effect modifiers, and effect modifications that are important to public health.

Cite this article as: Kamangar F. Effect modification in epidemiology and medicine. *Arch Iran Med*. 2012; **15(9):** 575 – 582.

Introduction

Causal factors do not act uniformly under all circumstances, a phenomenon that is known as heterogeneity of effect, effect modification, or interaction. In this article, we use these terms interchangeably. The following two examples show heterogeneity of effect.

Example 1: The antibiotic tetracycline could discolor teeth.[1] However, this effect depends on age. Tooth discoloration mainly happens in children under the age of eight.[1] In this case, we say that the effect of tetracycline (the exposure) on tooth color (the outcome) is modified by age (the effect modifier).●

Example 2: Abacavir, a drug used to treat HIV-infected patients, may cause a severe hypersensitivity that could even be fatal.[2] However, Abacavir (the exposure) causes hypersensitivity (the outcome) almost exclusively in those who have a certain allele, HLA-B*5701.[2] Therefore, the effect of the drug is modified by genetic alleles.●

Effect modification is present when the causal effect of an exposure on an outcome "depends" on a third factor. If one asks the question: "Does tetracycline discolor teeth?" The answer is "it depends on age". Likewise, if one asks whether Abacavir causes hypersensitivity, the answer is "it depends on whether the person carries the allele HLA-B*5701".

It is important to know how effect modification (interaction) is defined, identified, and interpreted. Therefore, in this article, we discuss the following topics:

1. Definition of effect modification;
2. Types of effect modification;
3. Methods used to identify effect modifiers;
4. Reasons for observing statistical interactions;
5. Statistical versus biologic interactions;
6. Special forms of interactions (gene-gene and gene-environment interactions);
7. Sample size required to detect statistical interactions;
8. Dependence of effect of modification on choice of the model;
9. Types of departure from additive and multiplicative models;
10. Choice of regression model and effect modification;
11. Bidirectionality of effect modification;
12. Factorial designs and interactions;
13. Interactions of significance to public health;
14. Effect modifiers versus other "third variables";
15. Interactions and adjustment.

1. Definition

If the effect of an exposure (E) on an outcome (O) depends on a third variable (M), M is an effect modifier; the effect of E on O is modified by M. In Example 1, the effect of tetracycline on tooth discoloration depended on age.

Two important points needs to be emphasized here. First, the term "effect" in the definition above in fact means a statistical association, which may or may not be causal (see Section 4). Second, this definition of effect modification may depend on what measure of statistical association (e.g., relative risk, odds ratio, or attributable risk) is used (see Sections 8 and 9).

2. Types of effect modification

Effect modification can be classified into qualitative effect modification and quantitative effect modification (Figure 1).

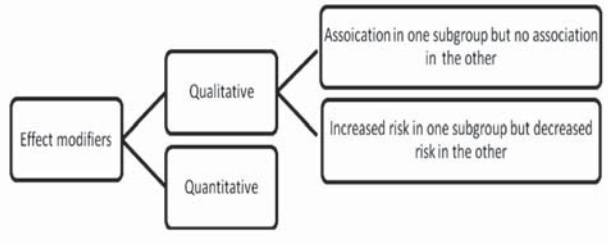

Figure 1. Qualitative versus quantitative effect modifiers

Qualitative effect modification is present when the direction (or nature or quality) of the effect depends on the effect modifier. Examples 1 and 2 both show qualitative effect modification. In Ex-

Authors' affiliations: [1]School of Community Health and Policy, Morgan State University, Baltimore, USA, [2]Digestive Disease Research Center, Tehran University of Medical Sciences, Tehran, Iran
•**Corresponding author and reprints:** Farin Kamangar MD PhD, Department of Public Health Analysis, School of Community Health and Policy, Morgan State University, Portage Avenue Campus, Room 302, Baltimore, MD 21251.
E-mail: farin.kamangar@morgan.edu
Accepted for publication: 10 July 2012

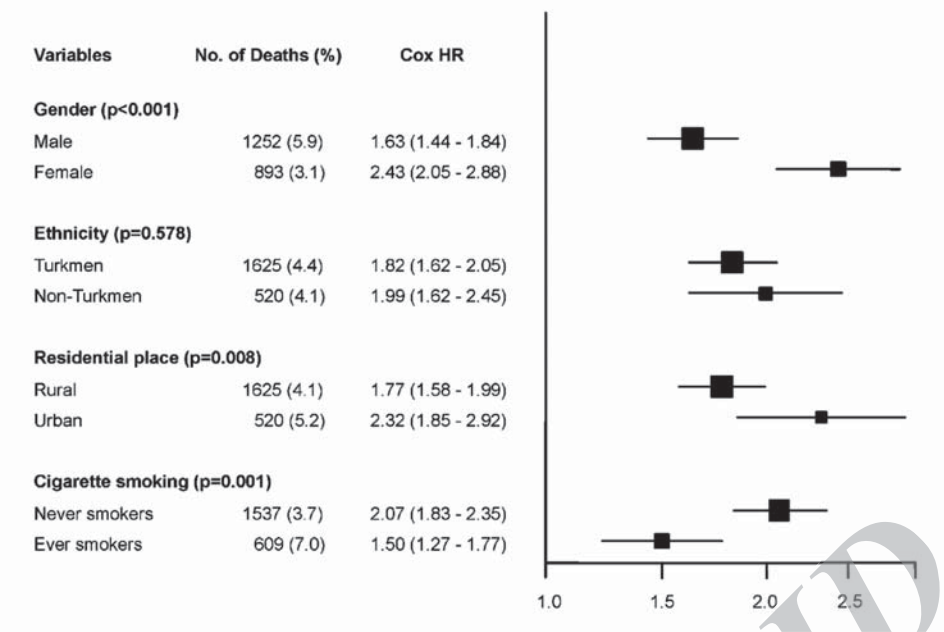

Figure 2. The effect of opium on mortality stratified by gender, ethnicity, residential place, and cigarette smoking.

ample 1, tetracycline increases the risk of tooth discoloration in younger individuals (relative risk > 1) but does not have an effect in older ages (relative risk = 1). Also, if the exposure increases the risk of the health outcome in some situations (relative risk > 1) and decreases the risk in other situations (relative risk < 1), that is considered qualitative effect modification.

Quantitative effect modification exists when the nature or direction of the association doesn't change but the strength (quantity of association) may differ across strata of the effect modifier. See Example 3.

Example 3: The results of a cohort study in Iran showed that after adjustment for several factors, opium increased total mortality risk by 86% (relative risk = 1.86).[3] When the researchers stratified the results by sex, this association was stronger for women (relative risk = 2.43) than for men (relative risk = 1.63); p-value < 0.001. In this example, opium increases the risk of death in both men and women so there is no qualitative difference. However, the effect of opium on death (in terms of relative risk) is stronger in women than in men. This is an example of quantitative effect modification.●

Please note that the difference across strata should be statistically significant at some alpha level (for example 0.05 or 0.10), otherwise differences in effect estimates may be considered to be due to chance.

Example 4: In the study of opium and mortality,[3] when results were stratified by ethnic group of the study participants, the relative risk of death associated with opium use in Turkmen and non-Turkmen ethnic groups were 1.82 and 1.99, respectively. However, p-value for interaction was 0.58, which was not statistically significant. Therefore, the authors concluded there was no effect modification (interaction) by ethnicity.●

3. Methods used to identify effect modifiers

Effect modifiers are usually identified using either stratification or regression methods.

3.1 Stratification

Stratification has already been discussed in the previous examples. In Example 3, the results were stratified by sex, and there was a statistically significant difference between men and women. In Example 4, the results were stratified by ethnicity, and there was no statistically significant difference. Stratified results are often tabulated or are represented graphically using forest plots. Figure 2 shows the forest plot for stratified results of opium use on mortality by gender, ethnicity, residential place, and cigarette smoking.

3.2 Regression

Regression models estimate the magnitude and statistical significance of effect modification by putting interaction terms in the models. Interaction terms are often constructed using product terms of the main exposure and the effect modifier. Further details on the algebra of interactions terms are provided in Box 1, Box 3, and Section 10.

Please note that stratification and regression are statistical methods. Therefore, they tell us about "statistical interactions". As mentioned in Section 1 and detailed in the next two sections, these interactions are a form of association and may or may not be causal.

4. Reasons for observing statistical interactions

Statistical interactions may arise due to a number of reasons including random variation (chance findings), differential confounding, differential bias, differential misclassification, differential intensity of exposure, competing risk factors, and biologic reasons.[4] The term "differential" here refers to a difference across strata of the effect modifier variable.

Interactions due to random variation are very commonly seen in epidemiologic studies. When researchers do not find overall significant results, they often conduct subgroup analyses, which may result in statistically significant results in one subgroup. However, many such findings are chance findings that result from increased type I error. Therefore, statisticians warn against immediate inter-

Box 1. Interaction terms in regression models

Interaction terms can be used in regression models to quantify the magnitude of interaction, and to test for statistical significance of interaction. Interaction terms are often the product of the exposure and the potential effect modifier. These terms can be interpreted most easily when both the exposure and the effect modifier are binary variables.

Here, we illustrate the simplest case where the regression model is linear and the exposure and the effect modifier are binary (0, 1) variables. The outcome, the exposure, and the effect modifier are named Y, X_1, and X_2. The model is shown here: $E(Y) = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_1 X_2$. In this model E(Y) is average of Y, b_0 is the intercept, $b_1 X_1$ is the main effect of X_1, $b_2 X_2$ is the main effect of X_2, and $b_3 X_1 X_2$ is the interaction term. Below, we show how this interaction term affects the results.

First, consider a situation that the effect modifier is absent ($X_2 = 0$). In this situation, in the absence of exposure ($X_1 = 0$), E(Y) will be b_0. In the presence of the exposure ($X_1 = 1$), E(Y) will be $b_0 + b_1$. Therefore, the presence of exposure changes E(Y), which is the average Y, by b_1.

Second, consider a situation that the effect modifier is present ($X_2 = 1$). In this situation, in the absence of exposure ($X_1 = 0$), E(Y) will be $b_0 + b_2$. In the presence of the exposure ($X_1 = 1$), E(Y) will be $b_0 + b_1 + b_2 + b_3$. Therefore, the presence of exposure changes E(Y), which is the average Y, by $b_1 + b_3$.

In summary, when the effect modifier is absent, exposure increases average Y by b_1, while when the effect modifier is present, the exposure increases average Y by $b_1 + b_3$. Therefore, the magnitude of effect of the exposure "depends" on the presence of effect modifier. Here, b_3, the coefficient of the interaction term, is the difference of the effect between when the modifier is present versus when it is not present.

pretation of subgroup analyses. Unless the findings are repeated in other studies, often the results of subgroup analyses are not taken seriously.

Differential confounding can result in finding statistical interactions. This happens when the association between the exposure and the outcome is confounded in one subgroup but not in the other. For example, assume that in the 1950s men drank coffee at work when they smoked, whereas women drank their coffee at home and did not smoke. In such a circumstance, drinking coffee could be associated with lung cancer in men but not in women. Effect modification by sex is the results of confounding, not any real biologic reason.

Different forms of bias that are limited to one study subgroup may be another reason for finding statistical interactions. For example, assume that women respond truthfully to questions about a certain exposure while men do not. This may result in seeing an association in one sex but not the other.

Likewise, differential measurement error may result in association in one subgroup but not the other. For example, assume that women can accurately report their intake of fruits and vegetables but men report their intake with substantial random error. If in reality intake of fruits and vegetables is associated with a 50% reduction in risk of lung cancer, that reduction will be detectable in women but not in men.

Differential intensity in exposure may lead to finding statistical interactions. For example, if men smoke more heavily than women and the results are not accurately adjusted for intensity of smoking, researchers may find that smoking is more strongly associated with lung cancer risk in men than in women.

Competing risk factors may be an important reason for finding statistical interactions. For example, researchers may find that passive smoking is a risk factor for lung cancer in never smokers but not in heavy smokers. The reason is that the large majority of lung cancers in heavy smokers are caused by their active smoking and in these individuals perhaps very little risk is added by a weak risk factor such as passive smoking. However, in never-smokers the effect of passive smoking may be more evident. Many such examples are found in epidemiology. A study of radon exposure and lung cancer found that radon is a stronger risk factor (in terms of relative risk) in never-smokers than in smokers.[5] Likewise, compared to low-risk areas of the world for esophageal squamous cell carcinoma, in high-risk areas tobacco smoking is a much less strong risk factor for this cancer.[6,7] The reason may be the presence of a very strong risk factor for esophageal squamous cell carcinoma that competes with smoking for causing this cancer.[8] Jerome Cornfield, a distinguished statistician who developed multiple logistic regression and found that odds ratios could be used as an approximation to relative risks in many case-control studies,[9] was one of the first statisticians who alluded to the effect of competing risk factors in causing interactions.[10]

Finally, and perhaps of most interest, biologic interactions could be the reason for finding statistical interactions. Several examples have been provided before or later in this paper, including the interaction between tetracycline and age in causing tooth discoloration,[1] Abacavir and HLA type in causing hypersensitivity reactions,[2] cheese intake and monoamine oxidases in causing hypertensive crisis,[11,12] and ADH2 and ALDH2 polymorphisms and alcohol intake in causing hypersensitivity to alcohol.[13,14]

5. Statistical interactions versus biologic interactions

The term "interaction" may not mean exactly the same thing to statisticians and biologists. As we will detail in Sections 8 to 10, statisticians consider interaction as departure from a model, such as from a linear (additive) or log-linear (multiplicative) model.[15] Statistical interaction is in fact a form of association and, as discussed in Section 4, may or may not have any causal implication. To biologists, however, interactions may mean that the causal effect of one variable is different in the presence of another one. For example, one protein may need another protein to accomplish a function. In the presence of the other one, the protein functions, and in the absence of the other one, it does not. Also, drug-drug interactions and food-drug interaction may be biologically detectable. For example, people who take monoamine oxidases (a class of antidepressants) should avoid or limit foods containing tyramine (such as aged cheese) to avoid risk of hypertensive crisis.[11,12] Neither eating usual quantities of cheese (e.g., 100 grams over one hour) nor taking monoamine oxidases alone causes hypertensive crisis, but when taken together they do cause such reactions.[16]

When epidemiologic studies show evidence for statistical interaction, many researchers may assume that this finding implies biologic interaction. Here, we will describe why this assumption may be incorrect. First, as mentioned in Section 4, there are many reasons to find a statistical association that are not causal. Second, as mentioned later (in Sections 8 and 9), statistical interaction depends on the scale of measurement. When there is no multiplicative interaction, there is additive interaction and vice versa. Therefore, there is always some form of interaction; it just depends on the scale one chooses. Rothman, Greenland, and Lash have shown that under certain assumptions – which are difficult to verify – departures from additivity imply biologic interaction. If so, most epidemiologic results that report no interaction may indeed be implying biologic interaction. This is because most of

the models that we use in epidemiology (e.g., logistic regression, Cox proportional hazards regression, Poisson regression) are multiplicative models, and when there is no multiplicative interaction, there is additive interaction. However, verifying the assumptions under which departure from additivity imply interaction is not always straightforward. Therefore, departures from additivity do not necessarily indicate biologic interactions. Third, lack of interaction, either additive or multiplicative, doesn't necessarily mean that there is no biologic interaction. This is because detection of statistical interaction often needs thousands of cases and controls, and many epidemiologic studies are underpowered to detect interactions. See Section 7.

6. Special forms of interaction

Some forms of interaction have been given names that are commonly used in epidemiological, biological, and medical literature. Examples are gene-gene interactions and gene-environment interactions. A PubMed search of "gene-environment interaction" on June 2, 2012, resulted in over 4,000 entries. As discussed earlier, these terms may have different meanings to biologists, physicians, and statisticians. Partly due to such differences in terminology and partly due to subject area, a biologist working on fetal development may find gene-gene interactions very common, whereas a statistician working on cancer susceptibility genes may find gene-gene interactions very uncommon.

To biologists and physicians, gene-environment interaction most often means that an environmental factor's effect on body depends on a genetic factor. For example, hypersensitivity to alcohol consumption in Asian populations depends on polymorphism in ADH2 and ALDH2 genes; those who carry certain polymorphisms are much more likely to develop signs and flushing and other symptoms of alcohol hypersensitivity.[13,14] Likewise, to biologists who work on fetal development, gene-gene interaction may mean that two or more genes are needed for a certain development to occur, and this is an extremely common phenomenon.

A statistician working on cancer genetics needs departures from a model (e.g., a log-linear model) to find an interaction. Because detecting such interactions requires a very large sample size, finding such interactions may be uncommon. While genome-wide association studies have found a large number of associations with cancer and other health outcomes, gene-gene interactions have been rarely found in such studies.

7. Sample size required to detect statistical interactions

Discussing the details of sample size calculation for interaction is beyond the scope of the article. However, there are at least two points worth mentioning. First, there are several free statistical software programs that can calculate sample size to detect statistical interaction, such as the Power V 3.0 program.[17,18] Second, as illustrated below, detecting interactions typically requires very large sample sizes.

Here we compare the sample size for two case-control studies, one with and one without interaction. The objective of the first case-control study is to detect an association between the exposure X_1 and the outcome Y. Assuming a case to control ratio of 1, two-sided type I error level of 0.05, power of 0.80, prevalence of exposure of 50% among controls, the study requires approximately 390 cases and 390 controls to detect an odds ratio of 1.5. The objective of the second case-control study is to determine the interaction between exposures X_1 and X_2 in causing outcome Y. Assuming that the prevalence of each exposure is 50% among controls and these two exposures are independently distributed, a case to control ratio of 1, two-sided type I error level of 0.05, power of 0.80, and that each of these exposures increases the risk of Y by 1.5-fold, the required sample size to detect an interaction of 1.5-fold would be 1690 cases and 1690 controls. While researchers might want to leave sample size calculations to statisticians, they should know that detecting statistical interactions typically require large sample sizes.

8. Dependence of effect modification on choice of the model

Earlier, in Section 1, we mentioned that the presence or absence of effect modification depends on the choice of the measure of association. To clarify this further, here we first introduce two models for the association between the exposure and the outcome, multiplicative and additive and models. Then we discuss the effect of choice of model on statements regarding effect modification.

8.1 Expected results under a multiplicative model

Examine table 1, which shows the percentage of a disease D in relation to the exposure X and the potential effect modifier M. Assume that sample sizes are very large and therefore no random errors exist. If the model is multiplicative, what do we expect to see in the cell with question mark?

Table 1. Percentage of a disease D in relation to the exposure X and the potential effect modifier M

	No X	X
No M	1%	4%
M	5%	?

When M is absent (No M row), X increases the risk by 4-fold. If X were going to do the same in the presence of M, then risk of disease D in the cell with question mark would be 20% (4 × 5%).

By definition, effect modification is present when the effect of X on D depends on M. If the number in the cell with question mark is 20%, the effect of X on disease D is ***multiplying*** its risk by 4-fold, regardless of the presence or absence of M. So, if that number is 20%, there is no effect modification when the effect is assumed to be multiplicative. However, if that number is not equal to 20%, then the effect of X on disease D depends on M. In the latter case (? ≠20%), we say there is "multiplicative effect modification" or "multiplicative interaction".

8.2 Expected results under an additive model

Table 2 is similar to what we saw in Section 4.1. However, here we ask what we expect if the model is additive.

Table 2. Percentage of a disease D in relation to the exposure X and the potential effect modifier M

	No X	X
No M	1%	4%
M	5%	?

In people not exposed to M, X increases the risk by 3%. If X were going to do the same in the presence of M, then risk of disease D in the presence of both X and Y would be 8% (3% + 5%).

If the number in the cell with question mark is 8%, the effect of X on disease D is ***adding*** to its risk by 3%, regardless of the presence or absence of M. So, if that number is 8%, there is no effect

modification when the effect is assumed to be additive. However, if that number is not equal to 8%, then the effect of X on disease D depends on M. In the latter case (? ≠ 8%), we say there is "additive effect modification" or "additive interaction".

8.3 Choice of model and effect modification

It should be clear by now that the choice of the model has a direct impact on statements regarding interaction. Assume that in the examples above the number in the cell with question mark is 20%. In such a situation, there is no multiplicative interaction, as X increases the risk by 4-fold, regardless of presence of M. However, there is there is additive interaction, as X adds to the risk by 3% in the absence of M and by 15% in the presence of M.

In fact, except for when the exposure has no effect, there is either multiplicative interaction or additive interaction or both, because the absence of one implies the presence of the other. So, the relevant question is not whether interaction exists; it is what type of interaction exists. See also Examples 5 and 6.

Example 5: Table 3 shows the percentage of the participants in a cohort study who were diagnosed with lung cancer according to their exposure to smoking and radon. Assume that these numbers come from a very large sample size, so the numbers given in the table are very precise. In this example, smoking is the exposure and lung cancer is the outcome. The question is whether radon is an effect modifier.

Table 3. Percentage of the participants in a cohort study who were diagnosed with lung cancer according to their exposure to smoking and radon

	Non-smokers	Smokers
Not exposed to Radon	1%	10%
Exposed to Radon	3%	30%

First, to examine multiplicative interaction, we assess the risk ratios. Among people who are not exposed to radon (the first row), smoking multiplies the risk of lung cancer by 10-fold (from 1% to 10%, or risk ratio = 10). Among people who are exposed to radon (the second row), again smoking multiplies the risk of lung cancer by 10-fold (from 3% to 30%, or risk ratio = 10). Therefore, risk ratios are similar across the strata of radon, or the effect of smoking on lung cancer in terms of risk ratios *does not depend* on radon. Statistically speaking, radon does not modify the effect of smoking on lung cancer in terms of risk ratios, or there is no multiplicative interaction between radon and smoking in causing lung cancer.

Second, to examine additive interaction, we assess risk differences. Among people who are not exposed to radon, smoking adds to the risk of lung cancer by 9% (from 1% to 10%), whereas among people who are exposed to radon, smoking adds to the risk of lung cancer by 27% (from 3% to 30%). Therefore, risk differences are different across the strata of radon, or the effect of smoking on lung cancer in terms of risk differences *does depend* on radon. In statistical terms, radon modifies the effect of smoking on lung cancer in terms of risk differences, or there is an additive interaction between radon and smoking in causing lung cancer.●

Example 6: Table 4 shows the percentage of the participants in a randomized 2x2 factorial trial of two drugs X and Y to treat an infectious disease. Assume that these numbers are very precise. In this example, X is the exposure and the infectious disease is the outcome. We would like to know whether Y modifies the effect of X on the disease.

Table 4. Percentage of the participants in a randomized 2x2 factorial trial of two drugs X and Y to treat an infectious disease

	No X	X
No Y	20%	40%
Y	40%	60%

First, we assess the risk ratios. Of those who receive neither treatment (the reference group), 20% recover spontaneously. Among those who do not receive Y (first row), X doubles the chance of recovery (from 20% to 40%). Among those who do receive Y, X increases the chance of recovery by 1.5-fold. These two are not the same, therefore Y modifies the multiplicative effect of X on disease (changes the risk ratio from 2 to 1.5). We may also say there is multiplicative interaction.

Now, we assess differences. In those who do not receive Y, X adds to the chance of recovery by 20%. Among those who do receive Y, again X increases the chance of recovery by 20%. Therefore, additions do not depend on status of Y. In other words, there is no additive interaction. In this example, unlike the previous one, there was no additive interaction but there was multiplicative interaction. This emphasizes again that when one type of interaction is absent, the other one is present.●

9. Types of departure from multiplicative and additive models

When the results depart from either multiplicative or additive models, they may be more than the expectation or less than the expectation. In such cases, terms that are used are different.

When the results are beyond our expectation from a multiplicative model, we call it positive multiplicative interaction. In the example used in Section 9, if the number in the cell with question mark is more than 20%, there is positive multiplicative interaction. If the number is less than 20%, there is negative multiplicative interaction. If the number is 20%, there is no multiplicative interaction, as there is no departure from multiplicative model.

Likewise, when the results are beyond expectation from an additive model, we call it positive additive interaction. In the example used in Section 9, if the number in the cell with question mark is more than 8%, there is positive additive interaction. If the number is less than 8%, there is negative additive interaction.

The arrow shown in Figure 3 shows positive and negative interactions based on multiplicative and additive models for the examples in Section 9. In this graph, any interaction that is positive multiplicative, is also positive additive; and any interaction that is negative additive, is also negative multiplicative. But there may be cases where there is positive additive but negative multiplicative interaction. This is a general case. For an algebraic proof, see Box 2.

Figure 3. In this figure, the white area (< 8%) shows negative additive and negative multiplicative interaction. The grey area (> 8% but < 20%) shows positive additive but negative multiplicative interaction. The black area shows positive additive and positive multiplicative interaction.

10. Choice of regression model and effect modification

Box 2. The multiplicative model expectation is always beyond the additive model expectation

Assume that X_1 increases the risk of Y by e_1, therefore the relative risk associated with X1 is $(1+e_1)$. X2 increases the risk of Y by e_2, or the relative risk associated with X2 is $(1+e_2)$.
Under an additive model, if both X_1 and X_2 are present, we expect that the risk of Y will be $1+e_1+e_2$. Under a multiplicative model, when both X_1 and X_2 are present, we expect that the risk of Y will be $(1+e_1) \times (1+e_2) = 1 + e_1 + e_2 + e_1 \times e_2$. Expectation under the multiplicative model is more than the additive model by $e_1 \times e_2$. Therefore, positive multiplicative interaction always implies positive additive interaction, and negative additive interaction always implies negative multiplicative interaction. But there can be situations where there is positive additive but negative multiplicative interaction.

In this section, we discuss how the choice of the four most commonly regression models in epidemiology – i.e., logistic regression, Cox proportional hazards regression, Poisson regression, and linear regression – affects our interpretation of effect modification.

Logistic regression, Cox proportional hazards regression, and Poisson regression determine odds ratios, and hazard ratios, and rate ratios respectively. Since these are all ratio measures – i.e., they show how many folds the risk (or odds or hazard or rate) is multiplied by in the presence of exposure – these models can determine multiplicative interactions. In contrast, linear regression which has a direct link to the outcome determines additive interaction, however, risk is rarely modeled using linear regression. Since logistic, Cox, and Poisson regression models are very popular is epidemiology to model the effect of a risk factor on an outcome, and all of these models estimate multiplicative interaction, the default meaning of interaction in epidemiology is "multiplicative interaction".

It is useful to know that interaction terms in multiplicative models compare the risk ratio (or odds ratio or hazard ratio) in the presence versus in the absence of the effect modifier. In Example 5, tobacco increases the risk of lung cancer in the presence of radon by 10-fold and in its absence by 10-fold. Therefore, the estimated interaction is 1 (10 ÷ 10 = 1), which indicates no multiplicative interaction. In Example 6, X increases the chance of treatment by 1.5-fold in the presence of Y and by 2-fold in the absence of Y. Therefore, the estimated interaction is 0.75 (1.5 ÷ 2 = 0.75), which shows sub-multiplicative interaction. If you are interested in an algebraic explanation, see Box 3.

11. Bidirectionality of effect modification

Thus far, we have studied the effect of a main exposure (X) on an outcome (O) in the presence or absence of an effect modifier (M). In epidemiologic studies, we are often interested in a certain exposure. Consider Example 1. The main exposure is tetracycline and age is only an effect modifier. However, this is not always the case. We may be interested in two different exposures and their interaction in causing the outcome. In situations like this, if the first exposure modifies the effect of the second exposure, the second exposure also modified the effect of the first one.

In Example 5, assume we are interested the effect of both smoking and radon and their interaction in causing lung cancer. In that example, smoking increases the risk of lung cancer by 10-fold regardless of whether the person is exposed to radon. So, there is no multiplicative interaction when we look at the results by row. Likewise, radon increases the risk of lung cancer by 3-fold, regardless of cigarette smoking status. So, there is no multiplicative interaction when we examine the results by columns either. In this same table, there is additive interaction, when we look at the results by rows (the effect of smoking) or by columns (the effect of radon).

In Example 6, again we see bidirectionality of interaction. Results show no additive interaction by rows (increase by 20% in each row). Results show no additive interaction by column either (increase by 20% in each column). However, examining the results by both rows and columns shows multiplicative interaction.

12. Factorial designs and interactions

Here we limit the discussion to 2×2 factorial designs, as they are the most common of factorial designs in epidemiology and medicine. In this design, two interventions, X and Y, are used to treat or prevent Disease D. The participants are randomized into four treatment arms: those who receive neither X nor Y; those who receive Y but not X; those who receive X but not Y; and those who receive both. Table 5 shows the design.

Table 5. 2×2 Factorial design trials

	No X	X
No Y	Disease D %	Disease D %
Y	Disease D %	Disease D %

From the discussions above, it should be clear that factorial designs allow for evaluating the effect of X on Disease D, the effect of Y on Disease D, and the interaction between the X and Y. A classic example is the Alpha-Tocopherol, Beta-Carotene Cancer

Box 3. Interaction terms in Cox proportional hazards regression and other multiplicative models

Here, we illustrate a simple case where the exposure and the effect modifier are binary (0, 1) variables. The outcome, the exposure, and the effect modifier are named Y, X_1, and X_2.
In Cox regression models, the natural logarithm of hazard of the outcome is modeled. The model is shown here: $E(\log \text{hazard of Y}) = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_1 X_2$. In this model b_0 is the intercept, $b_1 X_1$ is the main effect of X_1, $b_2 X_2$ is the main effect of X_2, and $b_3 X_1 X_2$ is the interaction term. Below, we show how this interaction term affects the results.
First, consider a situation that the effect modifier is absent ($X_2 = 0$). In this situation, in the absence of the exposure ($X_1 = 0$), E(log hazard of Y) will be b_0, or the hazard of Y will be $\exp(b_0)$. In the presence of the exposure ($X_1 = 1$), E(log hazard of Y) will be $b_0 + b_1$ or the hazard of Y will be $\exp(b_0+b_1) = \exp(b_0) \times \exp(b_1)$. Therefore, the presence of the exposure multiplies the hazard by $\exp(b_1)$. In other words, the hazard ratio is $\exp(b_1)$.
Second, consider a situation that the effect modifier is present ($X_2 = 1$). In this situation, in the absence of the exposure ($X_1 = 0$), E(log hazard of Y) will be $b_0 + b_2$, or the hazard of Y will be $\exp(b_0+b_2) = \exp(b_0) \times \exp(b_2)$. In the presence of the exposure ($X_1 = 1$), E(log hazard of Y) will be $b_0 + b_1 + b_2 + b_3 = \exp(b_0) \times \exp(b_1) \times \exp(b_2) \times \exp(b_3)$. Therefore, the presence of exposure multiplies the hazard by $\exp(b_1) \times \exp(b_3)$. In other words, the hazard ratio is $\exp(b_1) \times \exp(b_3)$.
In summary, when the effect modifier is absent, the hazard ratio of the outcome associated with exposure X_1 is $\exp(b_1)$, whereas when the effect modifier is present, the hazard ratio is $\exp(b_1) \times \exp(b_3)$. Therefore, the magnitude of effect of the exposure "depends" on the presence of effect modifier. Here, $\exp(b_3)$, the coefficient of the interaction term, is the ratio of the effect comparing when the modifier is present versus when it is not present.

Prevention Study (ATBC Study), in which participants were randomized into receiving alpha-tocopherol only, beta-carotene only, both, or neither.[19] The aim of this study was to determine whether any of these compounds, or a combination of them, can reduce the risk of lung cancer. Many other well-known examples are found in the medical literature.[20,21]

13. Interactions of significance to public health

Some authors have argued that if two exposure cause a disease beyond an additive model (positive additive interaction), that should be considered of significance to public health.[22] An example will clarify why this is the case.

Assume that the lifetime risk of a disease in the absence of its two main causal risk factors (A and B) is 1%. Again, assume that risk factor A alone increases the risk by 5% (increases by 4%, or 5-fold) and risk factor B alone increases the risk from 1% to 6% (by 5%, or 6-fold). If both risk factors are present and the model is additive, they should increase the risk to 10% (1% + 4% + 5%). In such a condition, removing factor A decreases the risk by 4%, regardless of the presence of B. However, if the model is multiplicative, which is beyond additive, the risk of this disease in the presence of both risk factors will be 30% (1% × 5 × 6). In such a condition, removing A when B is present will decrease the risk the risk from 30% to 6% (by 24%). Therefore, in the presence of an interaction that is beyond additive, public health actions in reducing at least one of the risk factors become very important. Ironically, in most statistical models used in epidemiologic studies, interaction means departure from a multiplicative model, not an additive model.

14. Effect modifiers versus other third variables (confounders and mediators)

The exposure and the outcome are the two main variables when we examine an association. An effect modifier is a "third variable" that modifies the association between the exposure and the outcome. However, there are other third variables that may affect the association between the exposure and the outcome, including confounders and mediators. The distinction between effect modifiers, confounders, and mediators has been described in a previous article.[23]

It is important to note that mediators in general are effect modifiers. For example, assume that in some countries poverty (the exposure) increases the risk of low birth weight (the outcome) because of mother's limited access to adequate and nutritious food (the mediator). If the government of a country decides to provide adequate and nutritious food to all mothers, then poverty will not increase the risk of low birth weight. Therefore, provision of adequate and nutritious food modifies the effect of exposure on the outcome. However, the converse is not necessarily true. As discussed in Section 4, effect modifiers are not necessarily mediators.

15. Interactions and adjustment

When an exposure acts as a confounder, we usually adjust for it to find the correct estimate for the causal association between the exposure and the outcome. However, do we adjust when the exposure is both a confounder and an effect modifier?

There are occasions where an exposure may be both a confounder and an effect modifier.[23] For example, the unadjusted odds ratio for the association between exposure E and disease D may be 2.00. When we stratify by sex, the odds ratio for men and women are 1.50 and 1.00 (p for interaction = 0.001), respectively, and the Mantel-Haenszel weighted average of these two numbers is 1.20. Here, sex acts as a confounder, as the adjusted odds ratio of 1.20 is different from the unadjusted odds ratio of 2.00. Sex also acts as an effect modifier, as the odds ratio is different for men and for women.

The decision to adjust depends on whether or not we believe the differences are due to chance; whether the differences are qualitative or quantitative; and how different the relative risk estimates are when we stratify them. When we stratify by subgroups, each subgroup may have a small number of subjects and therefore widely different relative risks may be found among the subgroups solely due to chance. If so, we tend to take a weighted average of the results and report adjusted relative risks. If, however, the relative risk estimates are statistically significantly different when stratified and we believe this is not due to chance, we may report a combined adjusted value, or separate values, based on the magnitude and type of differences. For example, if the relative risk is 1.90 for men and 2.15 for women (p = 0.001), even though the results are statistically different across the strata of sex, we may choose to combine the results (e.g., relative risk =2.03), as they both indicate an almost an almost doubling of the risk in men and women. In contrast, if the relative risk is 3.00 for men but 0.50 for women (p = 0.001), we may choose to report the results separately, not combined, as the exposure increases the risk of the disease of interest in men but decreases the risk in women. In this latter case, a combined relative risk (e.g., relative risk =1.50) may conceal the different effect of the exposure on men and women.

Conclusions

Effect modification, also called interaction or heterogeneity of effect, occurs when the effect of an exposure on an outcome depends on a third variable, called the effect modifier. However, the presence of interaction or lack thereof depend on what the term "effect" means, or on the choice of model. The model may be additive or multiplicative; no additive interaction implies multiplicative interaction, and vice versa. Effect modification can be statistically detected using stratification or regression methods. While using such statistical techniques have sometimes led to finding important biological facts, we need to keep in mind that statistical interactions do not necessarily imply biological interactions; random variation, confounding, bias, measurement error, presence of competing risk factors, and differential intensity of exposure are among other reasons that can lead to statistical interactions. Alternative explanations should be ruled out before making a biologic conclusion. Finally, results that are beyond expectations from an additive model may have significant public health implications.

Acknowledgements

The author thanks Dr. Farhad Islami (Mount Sinai School of Medicine, New York, NY), Dr. Mahsa Mohebtash (Union Memorial Hospital, Baltimore, MD), Dr. Ramin Shakeri (National Cancer Institute, Rockville, MD), and Dr. Payam Sheikhattari (School of Community Health and Policy, Morgan State University, Baltimore, MD) for reading the paper thoroughly and providing constructive comments. The author also thanks Dr. Hooman Khademi (International Agency for Research on Cancer, Lyon, France) for providing Figure 1.

References

1. Tredwin CJ, Scully C, Bagan-Sebastian JV. Drug-induced disorders of teeth. *J Dent Res*. 2005; **84(7):**596 – 602.
2. Ma JD, Lee KC, Kuo GM. HLA-B*5701 testing to predict abacavir hypersensitivity. *PLoS Curr*. 2010; **2:**RRN1203.
3. Khademi H, Malekzadeh R, Pourshams A,Jafari E, Salahi R, Semnai S, et al. Opium use and mortality in Golestan Cohort Study: prospective cohort study of 50 000 adults in Iran. *BMJ*. 2012; **344:**2502. doi: 10.1136/bmj.e2502.
4. Szklo M, Nieto FJ. Epidemiology: Beyond the Basics. 2nd eds. Jones and Bartlett Publishers, Sudbury, MA 2007.
5. Lubin JH, Boice JD,Edling C,Hornung RW, Howe GR, Kunz E, et al. Lung cancer in radon-exposed miners and estimation of risk from indoor exposure. *J Natl Cancer Inst*. 1995; **87(11):**817 – 827.
6. Nasrollahzadeh D, Kamangar F, Aghcheli K,Sotoudeh M, Islami F, Abnet CC, et al. Opium, tobacco, and alcohol use in relation to oesophageal squamous cell carcinoma in a high-risk area of Iran. *Br J Cancer*. 2008; **98(11):**1857 – 1863.
7. Tran GD, Sun XD, Abnet CC,Fan JH, Dawsey SM, Dong ZW, et al. Prospective study of risk factors for esophageal and gastric cancers in the Linxian general population trial cohort in China. *Int J Cancer*. 2005; **113(3):**456 – 463.
8. Kamangar F, Malekzadeh R, Dawsey SM, Saidi F. Esophageal cancer in Northeastern Iran: a review. *Arch Iran Med*. 2007; **10(1):**70 – 82.
9. Greenhouse SW. Jerome Cornfield's contributions to epidemiology. *Biometrics*. 1982; **38:**33 – 45..
10. Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst*. 1959; **22(1):**173 – 203.
11. Horwitz D, Lovenberg W, Engelman K, Sjoerdsma A. Monoamine oxidase inhibitors, tyramine, and chesse. *JAMA*. 1964; **188:**1108 – 1110.
12. Asatoor AM, Levi AJ, Milne MD. Tranylcypromine and cheese. *Lancet*. 1963; **2(7310):**733 – 734.
13. Saito Y, Sasaki F, Tanaka I,Sato M, Okazawa M, Sakakibara H, et al. Acute severe alcohol-induced bronchial asthma. *Intern Med*. 2001; **40(7):**643 – 645.
14. Takeshita T, Mao XQ, Morimoto K. The contribution of polymorphism in the alcohol dehydrogenase beta subunit to alcohol sensitivity in a Japanese population. *Hum Genet*. 1996; **97(4):**409 – 413.
15. Cordell HJ. Detecting gene-gene interactions that underlie human diseases. *Nat Rev Genet*. 2009; **10(6):**392 – 404.
16. Wimbiscus M, Kostenko O, Malone D. MAO inhibitors: risks, benefits, and lore. *Cleve Clin J Med*. 2010; **77(12):**859 – 882.
17. The US National Cancer Institute Website.Available from: URL: http://dceg.cancer.gov/tools/design/POWER. (Accessed date: July 2012).
18. Garcia-Closas M, Lubin JH. Power and sample size calculations in case-control studies of gene-environment interactions: comments on different approaches. *Am J Epidemiol*. 1999; **149(8):**689 – 692.
19. The ATBC Cancer Prevention Study Group. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. The Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group. *N Engl J Med*. 1994; **330(15):**1029 – 1035.
20. Gaziano JM, Glynn RJ, Christen WG, Kurth T, Belanger C, MacFadyen J, et al. Vitamins E and C in the prevention of prostate and total cancer in men: the Physicians' Health Study II randomized controlled trial. *JAMA*. 2009; **301(1):**52 – 62.
21. Lippman SM, Klein EA, Goodman PJ, Lucia MS, Thompson IM, Ford LG, et al. Effect of selenium and vitamin E on risk of prostate cancer and other cancers: the Selenium and Vitamin E Cancer Prevention Trial (SELECT). *JAMA*. 2009; **301(1):**39 – 51.
22. Blot WJ, Day NE. Synergism and interaction: are they equivalent? *Am J Epidemiol*. 1979; **110(1):**99 – 100.
23. Kamangar F. Confounding variables in epidemiologic studies: basics and beyond. *Arch Iran Med*. 2012; **15(8):**508- 516.