

## Effect Modification - Part 3

### 1.1 Effect Modification



We will be formalizing what we have learned so far by introducing equations to examine interaction on additive and multiplicative scales.

### 1.2 What we will cover this unit:

#### What we will cover this unit:

1. Two definitions of effect modification (but they are really the same definition).
2. The use of interaction terms in a linear model.
3. The use of Interaction terms in a log-linear model
4. **Formally: scales of effect modification**
5. Introduction to RERI –relative excess risk due to interaction
6. How to test for effect modification
7. More on the Relative Excess Risk due to Interaction: Why is additive better than multiplicative?
8. More on interaction in a logistic model (computing correct odds ratios with an interaction term)
9. Does effect modification have any relationship to confounding (answer: Not really.)

### 1.3 To evaluate interaction between risk factors A and B

To evaluate interaction between risk factors A and B

	Risk Factor A	No Risk Factor A
Risk Factor B	R11	R01
No Risk Factor B	R10	R00

To evaluate where the interaction between risk factors A and B falls on an additive and on a multiplicative scale, you need to divide your population up into four groups of people.

Let's walk through who these four groups are, and the corresponding notation we will use. First, we have people with both risk factors, in the purple color of the two by two table. This risk of outcome in people in this group as R11. Then we have people with factor A and without B, these are R10. These are in the red color. Next are people without factor A and with factor B. These people are denoted R01 and blue. Finally, we have the people without both risk factors denoted as R00.

### 1.4 Interaction is Additive

Interaction is **Additive**

	Risk Factor A	No Risk Factor A
Risk Factor B	R11	R01
No Risk Factor B	R10	R00

If  $R11 - R00 = (R10 - R00) + (R01 - R00)$

Now here is a useful algebraic equality

$$R11 - R00 = (R10 - R00) + (R01 - R00) \quad RR11 = RR10 + RR01 - 1$$

Where  $RR11 = R11/R00$   
 $RR10 = R10/R00$   
 $RR01 = R01/R00$

Using the 4 groups of people defined on the previous slide, we use the definition of additive interaction written in notation, to derive a formula that allows us to evaluate additive interaction using risk ratios. This is very useful to us because as epidemiologists we tend to use ratios. Click on the text about algebraic equality to learn more about this topic.

### Important Algebraic Equality

- $R_{11} - R_{00} = (R_{10} - R_{00}) + (R_{01} - R_{00})$   
*Add  $R_{00}$  to both sides*
- $R_{11} = R_{10} + R_{01} - R_{00}$   
*Divide both sides by  $R_{00}$*
- $R_{11}/R_{00} = R_{10}/R_{00} + R_{01}/R_{00} - R_{00}/R_{00}$

$$RR_{11} = RR_{10} + RR_{01} - 1$$

Return

{ this is a layer of the previous slide } So you see all you do to get from the definition of additive interaction to a formula using risk ratios (which are easier for us to use) is to add  $R_{00}$  to both sides and then divide both sides by  $R_{00}$ .

## 1.5 Evaluating Additivity

### Evaluating Additivity

If (actual)  
 $RR_{11} > RR_{10} + RR_{01} - 1$   
then it's **greater** than  
additive

**Synergistic**

If (actual)  
 $RR_{11} < RR_{10} + RR_{01} - 1$   
then it's **less** than additive

**Antagonistic**



To determine where the interaction falls on the additive scale, we compare the relative risk when both factors are present to the sum of the relative risk when only one factor is present minus 1. Take a moment to study this slide. You will see these inequalities used in a slightly different manner later. Also note that \*chance plays a role in the two sides of the equation being different. In the absence of a statistical test, you will have to use your best judgement as to whether the combination of factors is synergistic. For example a 2.3 and a 2.6 are probably not different, but a 2.3 and a 5.6 are.

However, remember that there is chance involved, because you are dealing with different subgroups. Don't pronounce two numbers different based on the eyeball test, unless they are very different.

## 1.6 Interaction is Multiplicative

### Interaction is *Multiplicative*

	Risk Factor A	No Risk Factor A
Risk Factor B	R11	R01
No Risk Factor B	R10	R00

$$RR_{11} = RR_{10} * RR_{01}$$

- The risk ratio for A and the outcome does not depend on B
- and
- The risk ratio for B and the outcome does not depend on A

Using the notation, here is the equation for multiplicative interaction. For multiplicative interaction the risk ratio for A and the outcome does not depend on B and the risk ratio for B and the outcome does not depend on A. Click on the equation to learn more about this notation.

### Notation: Risk Difference

	Risk Factor A	No Risk Factor A
Risk Factor B	R11	R01
No Risk Factor B	R10	R00

Risk Difference	
$RD_{11} = R_{11} - R_{00}$	Risk for the group with both factors minus the risk for the group with neither factor.
$RD_{10} = R_{10} - R_{00}$	Risk for the group with factor A only minus the risk for the group with neither factor.
$RD_{01} = R_{01} - R_{00}$	Risk for the group with factor B only minus the risk for the group with neither factor.

Next

We have a few more equalities about interaction to show you, and they require using notation. You do not have to use this particular notation. You may prefer to use capital and small letters for subscripts, as long as you know what it means. But let's recall what we discussed on the last slide.

Have two factors A and B. Use

- $R_{11}$  = Risk of outcome in group with both factors.
- $R_{10}$  = Risk of outcome in group with factor A, but not factor B
- $R_{01}$  = Risk of outcome in group with factor B, but not factor A
- $R_{00}$  = Risk of outcome in group with neither factor ("baseline risk")

We can use the risk difference to understand the risk difference for both factors, the risk for factor A only, and the risk for factor B only. Take some time to review this, then move on to the next page.

### Notation: Risk Ratio

	Risk Factor A	No Risk Factor A
Risk Factor B	R11	R01
No Risk Factor B	R10	R00

Risk Ratio	
$RR_{11} = R_{11} / R_{00}$	Risk ratio of the group with both factors relative the group with neither factor.
$RR_{10} = R_{10} / R_{00}$	Risk ratio of the group with factor A relative the group with neither factor.
$RR_{01} = R_{01} / R_{00}$	Risk ratio of the group with factor B relative the group with neither factor.

Return

Similarly, we can examine the risk ratio for each of these factors. But this time we are dividing rather than subtracting. Again, take some time to review this table. Once you're done click on the return button.

## 1.7 Two Definitions of Additive Interaction

### Two Definitions of Additive Interaction

There are two definitions for two factors combining additively.  
Let E1 and E2 be two exposures and D denote disease.

#1: *difference* measures of E1,D association are equal in all strata of E2 and vice versa

#2: *the difference measure for the combined effect of E1 and E2 is the sum of the difference measures for E1 only and for E2 only.*

PROOF

Now that we have reviewed the two scales of interaction, additive and multiplicative, let's go into some more detail. For the additive scale, recall there are two definitions of additive interaction. Let E1 and e2 be two exposures and D denote disease. The first definition is the *difference* measures of E1,D association are equal in all strata of E2 and vice versa. The second is the difference measure for the combined effect of E1 and E2 is the sum of the difference measures for E1 only and for E2 only. But fact, 1. and 2. are the same definition. We can use the notation to easily prove this

As a reminder, there are two definitions of additive interaction and now we are ready to prove that the two definitions are the same definition. Click on the proof button to learn more.



### Proof that Definitions #1 and #2 are the SAME

#### According to Definition #1 1

$$R_{10} - R_{00} = R_{11} - R_{01}$$

- the left side of eq is the effect of E1 for strata of E2 = 0
- the right side of eq is the effect of E1 for strata of E2 = 1

#### Add R01 to both sides

- $R_{11} = R_{10} + R_{01} - R_{00}$

#### Subtract (another) R00 from each side and rearrange

- $R_{11} - R_{00} = (R_{10} - R_{00}) + (R_{01} - R_{00})$
- $RD_{11} = RD_{10} + RD_{01}$       qed. 2

Return

So here is the proof that the definitions are the same. If you need to, use the markers to refresh on these definitions.

If you examine the proof, you will see that we go from definition 1 of additive interaction to definition 2 in three easy algebraic steps.

## 1.8 Two Definitions of Multiplicative Interaction

### Two Definitions of *Multiplicative Interaction*

There are two definitions for two factors combining additively.  
Let E1 and E2 be two exposures and D denote disease.

#1: ratio measures of E1, D association are equal in all strata of E2 and vice versa

#2: the ratio measure for the combined effect of E1 and E2 is the product of the ratio measures for E1 only and for E2 only.

PROOF

There are also two definitions of multiplicative interaction that are really the same definition. Note specifically that definition 1 is saying that relative risks (or odds ratios) for factor A and the outcome are the same for every strata of B and vice-versa. You probably recall this from Epi 6000. We leave the proof up to you. Take some time to try this out on your own before moving forward.

### Proof that Definitions #1 and #2 are the SAME

#1: ratio measures of E1, D association are equal in all strata of E2 and vice versa

#2: the ratio measure for the combined effect of E1 and E2 is the product of the ratio measures for E1 only and for E2 only.

**Define:**

$$RR_{AB} = R_{AB}/R_{ab}$$

$$RR_{Ab} = R_{Ab}/R_{ab}$$

$$RR_{aB} = R_{aB}/R_{ab}$$

$$RR_{B|A} = R_{AB}/R_{Ab}$$

$$RR_{B|a} = R_{aB}/R_{ab} (=RR_{aB})$$

$$RR_{A|B} = R_{AB}/R_{aB}$$

$$RR_{A|b} = R_{Ab}/R_{ab} (=RR_{Ab})$$

• Switching to notation, the problem becomes:

• Prove

$$RR_{B|A} = RR_{B|a} (=RR_{aB})$$

and

$$RR_{A|B} = RR_{A|b} (=RR_{Ab})$$

If and only if

$$RR_{AB} = (RR_{Ab})(RR_{aB})$$

Next

The trick to doing this exercise is to put each of the two conditions into notation. Once they are both in notation, the proof is trivial. Notation: Use capital letters if the factor is present and small letters if the factor is absent.

• Condition 1. Ratio measures of E1, D association are equal in all strata of E2 and vice versa

**If and only if**

• Condition 2 .the ratio measure for the combined effect of E1 and E2 is the product of the ratio measures for E1 only and for E2 only.

### Proof that Definitions #1 and #2 are the SAME

#1: ratio measures of E1, D association are equal in all strata of E2 and vice versa

#2: the ratio measure for the combined effect of E1 and E2 is the product of the ratio measures for E1 only and for E2 only.

**Start with**

$$\bullet RR_{AB} = (RR_{Ab})(RR_{aB})$$

**Divide both sides by  $RR_{Ab}$**

$$\bullet RR_{AB}/RR_{Ab} = RR_{aB}$$

$$\text{But } RR_{AB}/RR_{Ab} = R_{AB}/R_{Ab} = RR_{B|A}$$

$$\bullet RR_{B|A} = RR_{aB}$$

Return

In our notation this means:

## 1.9 What we will cover this unit:

### What we will cover this unit:

1. Two definitions of effect modification (but they are really the same definition).
2. The use of interaction terms in a linear model.
3. The use of Interaction terms in a log-linear model
4. Formally: scales of effect modification
5. Introduction to RERI -relative excess risk due to interaction
6. How to test for effect modification
7. More on the Relative Excess Risk due to Interaction: Why is additive better than multiplicative?
8. More on interaction in a logistic model (computing correct odds ratios with an interaction term)
9. Does effect modification have any relationship to confounding (answer: Not really.)

Now let's move on to talking about the fun acronym - RERI, which stands for the relative excess risk due to interaction.

## 1.10 The RERI

### The RERI

The relative excess risk due to interaction, *RERI*, is a statistic that comes from the definition of additive interaction.

$$RERI = RR_{11} + RR_{10} - 1$$

- ✓ Both factors need to increase the risk of the outcome.
- ✓ If RERI is significantly greater than 0, then we have synergism.

The relative excess risk due to interaction, RERI, is a statistic that comes from the definition of additive interaction. It is often used by epidemiologists.

A couple things to keep in mind about the RERI \*To use, both factors have to increase risk of outcome.

If RERI is significantly greater than zero we have synergism between the factors.

Unfortunately the calculation of the variance of RERI is outside the scope of the class.



## 1.11 What we will cover this unit:

### What we will cover this unit:

1. Two definitions of effect modification (but they are really the same definition).
2. The use of interaction terms in a linear model.
3. The use of Interaction terms in a log-linear model
4. Formally: scales of effect modification
5. Introduction to RERI –relative excess risk due to interaction
6. **How to test for effect modification**
7. More on the Relative Excess Risk due to Interaction: Why is additive better than multiplicative?
8. More on interaction in a logistic model (computing correct odds ratios with an interaction term)
9. Does effect modification have any relationship to confounding (answer: Not really.)

We have spent a lot of time going over various definitions of effect modification. How do we statistically test for it? When do we statistically test for this?

## 1.12 How do we test for effect modification?

### How do we **test** for effect modification?

**TYPE 1:** It's a secondary objective for you. You are just interested in heterogeneity of your Risk ratios or Odds ratios.

- a) Breslow-Day test or
- b) Look at the significance of the interaction term in logistic/log-linear/proportional hazards model
- c) Be very careful in your conclusion statements

**TYPE 2:** It's your primary objective.

- a) You will need to calculate and test the RERI statistic.
- b) Your conclusions will be straight forward.

There are two scenarios of how we would test for effect modification.

#### READ

The Breslow-Day test tests whether risk ratios or odds ratios are the same across strata; This is the same as testing for interaction on a multiplicative scale. SAS Proc Freq will put out the Breslow-Day test when doing an adjusted analysis. The significance of the coefficient of an interaction term in a multiplicative model also is a test of multiplicative interaction. However, you need to be aware that if you are truly interested in whether there is synergism between two factors, you need to calculate RERI.

### 1.13 What we will cover this unit:

#### What we will cover this unit:

1. Two definitions of effect modification (but they are really the same definition).
2. The use of interaction terms in a linear model.
3. The use of Interaction terms in a log-linear model
4. Formally: scales of effect modification
5. Introduction to RERI –relative excess risk due to interaction
6. How to test for effect modification
7. **More on the Relative Excess Risk due to Interaction: Why is additive better than multiplicative?**
8. More on interaction in a logistic model (computing correct odds ratios with an interaction term)
9. Does effect modification have any relationship to confounding (answer: Not really.)

Okay, we reviewed how to test for effect modification, but let's return to our favorite, RERI and answer the question, why is additive better than multiplicative?

### 1.14 Evaluate on additive scale

Why do we need to evaluate effect modification on an **additive** scale?

We have said several times in this lecture that interaction should really be examined on an additive rather than a multiplicative scale. We go through some examples now to try to give you an intuitive feeling as to why this may be true. In fact, we hope to show you that evaluating interaction on a multiplicative scale may lead to ridiculous conclusions.

## 1.15 Why is evaluation on an additive scale important? – Example 1

<p>Why is evaluation on an additive scale important? – Example 1</p> <p>Risk of premature <b>heart attack</b> (ages 45-64) in <b>men versus women</b>. Suppose the risk is 0.35% in men and 0.05% in women, meaning that men were 7 times more likely to have an premature heart attack.</p> <p>Now it has been reported that <b>diabetes</b> doubles the risk of MI in men but increases the risk of MI in women by 5 fold.</p> <ul style="list-style-type: none"><li>• RR for MI, <small>diabetic vs non diabetic</small>, Men = 2</li><li>• RR for MI, <small>diabetic vs non diabetic</small>, Women = 5</li></ul> <p>And suppose this difference in RR is statistically significant.</p> <p><b>Would you conclude that diabetes is significantly worse for women than for men?</b></p> <p>Next</p>	<p>We start with an example of MI in men and women ages 45-64. Suppose that in this age range being male is a strong risk factor as men are seven times more likely to have a heart attack than women.</p> <p>Suppose now what we are really interested in is the association between diabetes and MI in this middle aged age group. We decide to stratify on gender. We find that Risk ratio for diabetes and MI among men is 2.0, but among women it's 5.0 and the Breslow Day test (or interaction term in a logistic model) is highly significant. We conclude diabetes is worse for women than for men. What is the problem with that?</p>
<p>Why is evaluation on an additive scale important? – Example 1</p> <p>Risk of premature <b>heart attack</b> (ages 45-64) in <b>men versus women</b>. Risk = 0.35% in men and 0.05% in women</p> <p><b>Diabetes risk:</b></p> <p>RR for MI, <small>diabetic vs non diabetic</small>, Men = 2 RR for MI, <small>diabetic vs non diabetic</small>, Women = 5</p> <p><b>Risk of MI in a diabetic man = 0.7% = 0.35% × 2; (RD=0.35)</b> <b>Risk of MI in a diabetic woman = 0.25% = 0.05% × 5 (RD=0.2)</b></p> <p>Return</p>	<p>Here's the problem. If we take the baseline risk of MI in men and women and multiply it by the relative increase in risk with diabetes, the <b>Risk of MI in a diabetic man is 0.7%</b>. The Risk of MI in a diabetic woman is 0.25%. Do you really want to say that diabetes is significantly worse for women than for men when men get a greater increment in risk and a diabetic man is still almost 3 times as likely to have an MI than a diabetic woman????</p>

## 1.16 Why is evaluation on an additive scale important? – Example 2

Why is evaluation on an additive scale important? – Example 2

Lung cancer rates/100,000 PY

	Smokers	Non-Smokers
Exposed to Asbestos	648	58
Not Exposed	123	11

The combination of asbestos and smoking is multiplicative

PROOF #1: Equality of RR

- RR for asbestos in smokers =  $648/123 = 5.27$
- RR for asbestos in non-smokers =  $58/11 = 5.27$

PROOF #2: Equality of RR

- RR for smoking in asbestos people =  $648/58 = 11.34$
- RR for smoking in non-asbestos people =  $123/11 = 11.18$  (essentially equal)

PROOF #3: Equality of RR

- Show  $RR_{11} = RR_{10} * RR_{01}$
- $648/11 = 123/11 * 58/11$
- $58.9 = 11.18 * 5.27$
- $58.9 = 58.9$  qed

More

For example #2 let's go back to the asbestos and smoking example.

In example 2, we create a table such that asbestos and smoking is exactly multiplicative.

Note that I changed the smoking-asbestos cell to 648 for illustrative purposes.

The combination of asbestos and smoking is multiplicative (except for rounding in the table to whole #s) **and we will prove this not once but three times.**

Why is evaluation on an additive scale important? – Example 2

Lung cancer rates/100,000 PY

	Smokers	Non-Smokers
Exposed to Asbestos	648	58
Not Exposed	123	11

$$\ln(p) = \alpha + (\beta_1 * \text{smoking}) + (\beta_2 * \text{asbestos}) + (\beta_3 * \text{smoking} * \text{asbestos})$$

QUESTION: What is  $\beta_3$ ?

ANSWER: Zero

More

We evaluate the RR for asbestos in each smoking stratum, and the RR for smoking in each asbestos stratum. We find that the RR for asbestos doesn't depend on smoking and vice-versa

Here we show that the RR when both factors are present equals the product of the individual RRs

### Why is evaluation on an additive scale important? – Example 2

Lung cancer rates/100,000 PY

	Smokers	Non-Smokers
Exposed to Asbestos	648	58
Not Exposed	123	11

$$\ln(p) = \alpha + (\beta_1 * \text{smoking}) + (\beta_2 * \text{asbestos}) + (\beta_3 * \text{smoking} * \text{asbestos})$$

$\beta_3$  is the coefficient for interaction, which is the degree of deviation from what is expected based on the model. But based on a log linear model, multiplicative is what is expected, so there is no deviation.

Conclusion 1: There is no interaction between smoking and asbestos in the development of lung cancer?

Conclusion 2: The effect of asbestos is the same for smokers and non-smokers?

More

If we ran the following model

Now we run a log-linear model putting in an interaction term. What is the coefficient on the interaction term? The answer is zero, of course, because the interaction is exactly multiplicative.

### Why is evaluation on an additive scale important? – Example 2

Lung cancer rates/100,000 PY

	Smokers	Non-Smokers
Exposed to Asbestos	648	58
Not Exposed	123	11

$$\ln(p) = \alpha + (\beta_1 * \text{smoking}) + (\beta_2 * \text{asbestos}) + (\beta_3 * \text{smoking} * \text{asbestos})$$

**Conclusion 2: The effect of asbestos is the same for smokers and non-smokers?**

- We can make a sound case that asbestos is worse for smokers than non-smokers. After all, if 100,000 non-smokers were exposed to asbestos, 47 (58-11) more people out of 100,000 would develop lung cancer.
- On the other hand, if 100,000 smokers were exposed to asbestos, 525 (648-123) more people out of 100,000 would develop lung cancer.
- 525 patients use a lot more health resources and cost a lot more money than 47 patients.

Return

What can we then conclude about  $\beta_3$ ?

Conclusion 1 no: **Hahahaha. Only if you wanted to do standup comedy. Everyone would laugh at you for refuting a very well known synergistic effect.**

Based in the numbers in the table, I can make a sound case that asbestos is worse for smokers than non-smokers. If 100,000 non-smokers were exposed to asbestos, more people out of 100,000 would develop lung cancer. If 100,000 smokers were exposed to asbestos, more people out of 100,000 would develop lung cancer. So how can I get away with saying that asbestos has the same effect for smokers and non-smokers based on zero interaction term?



## 1.17 Targeting of interventions...

### Targeting of interventions...

- ✓ Needs to be based on impact.
- ✓ Recall the number of events due to an exposure = the **risk difference** times the **number exposed**.
- ✓ Therefore, conclusions as something being “worse” for a subgroup has to be based on **risk differences** rather than ratios.



Conclusions as to a risk factor being “worse” for one subgroup rather than another have to be based on risk differences rather than risk ratios.

## 1.18 Correctly Evaluate Interaction: Calculate RERI for Example 2

### Correctly Evaluate Interaction: Calculate RERI for Example 2

	Smokers	Non-Smokers
Exposed to Asbestos	648	58
Not Exposed	123	11

$$RERI = RR_{11} - RR_{10} - RR_{01} + 1$$

$$RERI = 58.9 - 11.18 - 5.27 + 1 = 43.45$$
$$43.45 > 0$$

Since RERI is substantially greater than 0, (for the purposes of this class) conclude synergism.

To correctly evaluate interaction, calculate RERI. In this smoking/asbestos example RERI comes out to 43, which is a lot bigger than 0. We conclude that asbestos and smoking are synergistic.

If RERI is substantially greater than 0, (for the purposes of this class) conclude synergism.

$$RERI = 58.9 - 11.18 - 5.27 + 1 = 43.45$$

Does  $43.45 = 0$ ? I guarantee you it does not, so you may conclude asbestos and smoking are synergistic.

**Give up your career in comedy.**

Unfortunately the test of  $RERI = 0$  is beyond the scope of this class. You will have to take Advanced Research Methods.

## 1.19 Take Home Message

### Take Home Message

- ✓ If you are looking at interaction on a **multiplicative model**:
  - ✓ be careful of your conclusions.
  - ✓ you can conclude that risk ratios or odds ratios are (not) homogenous, but you cannot say that the greater risk ratio indicates a greater effect.
  - ✓ you can conclude that the risk of two factors combine multiplicatively, less than multiplicative, or more than multiplicative.
- ✓ Beyond that, you need to stop and calculate the **RERI**.

If you are looking at interaction in a multiplicative model

The take-away message is be careful when examining interaction.