Interaction Analysis

Tyler J. VanderWeele
Departments of Epidemiology and Biostatistics
Harvard T.H. Chan School of Public Health

Plan of Presentation

Part I (today): Concepts of Interaction

(1) Motivations for Interactions
(2) Key Distinction #1: Additive vs. Multiplicative Interactions
(3) Statistical Interactions
(4) Key Distinction #2: Interaction vs. Effect Heterogeneity
(5) Presenting Interaction Analyses

Part II (next time): Motivations and Specific Interaction Analyses

Part I. Concepts of Interaction

Introductory Examples

Figueiredo et al. (2004) studied the effects of XRCC3-T241M polymorphisms and various environmental factors on breast cancer risk

For XRCC3-T241M using a case-control study they found the OR for breast cancer for the M/M genotype was 1.47 (CI: 1.00, 2.15) times that of the reference T/T or T/M genotype

However the effect varied by strata of alcohol consumption:

OR for breast cancer (by strata of alcohol consumption and XRCC3-T241M)

	No Alcohol	Alcohol
T/T or T/M	1.00	1.12 (0.81-1.54)
M/M	1.21 (0.70-2.09)	2.09 (1.16-3.78)

It seems as though XRCC3-T241M polymorphisms do not have much effect unless accompanied by alcohol consumption

This is an example of what we might call a gene-environment interaction

Why Do We Care?

To say that there is an interaction on some scale between two exposures is a relatively uninteresting statement

If both exposures have an effect on the outcome, then there must be an interaction on some scale (either additive, or multiplicative, or both) (Lash et al., 2021)

Thus to assert that there is an interaction on some scale is only to say that both exposures have an effect

If an interaction analysis is to be of interest, we should identify the reasons for carrying out such an analysis

Motivations for Interaction Analysis

(1) Selecting subgroups to treat when resources are limited

(2) Identifying subgroups in which treatment may be harmful

(3) Understanding mechanisms for the outcome

(4) Assessing the generalizability of an effect estimate

(5) When it is not possible to intervene on the exposure, interactions can help identify effect modifiers to intervene upon to eliminate exposure effects

(6) Assessing whether the effect is "bigger" for one subgroup vs. another

[Note: We might include a product term in a statistical model not because we are interested in "interaction" but simply to add model flexibility]

Notation

G: genetic factor of interest
E: environmental factor of interest
Y: outcome of interest

For simplicity we will assume that G and E are binary
i.e. E = 0 for the environmental exposure absent; E = 1 for present
i.e. G = 0 for low genetic risk; G = 1 for high genetic risk

We will use “G” and “E” but these can be any two exposures (e.g. behavioral, cultural, both environmental, etc.)

The ideas presented here apply more generally to exposures that are not binary

Confounding: We will assume for now that the effects of both exposures are unconfounded, but will return to issues of confounding and bias later

Additive vs. Multiplicative Interactions

How do we measure interaction? There are two major classes of numerical interaction measures

Additive interaction measures:
"How much bigger as a *difference* is the effect of two exposures together than the *sum* of the effects of each exposure separately?"

Multiplicative interaction measures:
"How much bigger as a *ratio* is the effect of two exposures together than the *product* of the effects of each exposure separately?"

Additive Interactions

Suppose we had the following data on risks from a cohort study:

	E=0	E=1
G=0	0.02	0.05
G=1	0.04	0.15

Let $p_{ij} = P(Y=1 \mid G=i, E=j)$.

A natural way to assess interactions is to measure the extent to which the effect of the two factors together exceeds the effects of each considered individually:

$$(p_{11} - p_{00}) - [(p_{10} - p_{00}) + (p_{01} - p_{00})] = p_{11} - p_{10} - p_{01} + p_{00}$$

This is sometimes referred to as a measure of interaction on the additive scale

Additive Interactions

Data:

	E=0	E=1
G=0	0.02	0.05
G=1	0.04	0.15

Additive measure of interaction: $p_{11} - p_{10} - p_{01} + p_{00}$

If $p_{11} - p_{10} - p_{01} + p_{00} > 0$ the interaction is said to be positive or "super-additive"

If $p_{11} - p_{10} - p_{01} + p_{00} < 0$ the interaction is said to be negative or "sub-additive"

Here we have:

$p_{11} - p_{10} - p_{01} + p_{00} = 0.15 - 0.04 - 0.05 + 0.02 = 0.08 > 0$

i.e. a positive interaction

Multiplicative Interactions

Data:

	E=0	E=1
G=0	0.02	0.05
G=1	0.04	0.15

As an alternative to assessing interactions on the additive scale using risks we might consider a multiplicative scale using relative risks:

Let $RR_{11} = p_{11}/p_{00} = 0.15/0.02 = 7.5$
Let $RR_{10} = p_{10}/p_{00} = 0.04/0.02 = 2$
Let $RR_{01} = p_{01}/p_{00} = 0.05/0.02 = 2.5$

A measure of multiplicative interaction for risk ratios is:
$RR_{11} / (RR_{10} \times RR_{01}) = 7.5 / (2 \times 2.5) = 7.5 / 5 = 1.5$
If the multiplicative interaction is > 1 it is positive, < 1 it is negative

Multiplicative Interactions

Data:

	E=0	E=1
G=0	0.02	0.05
G=1	0.04	0.15

Let $RR_{11} = p_{11}/p_{00} = 0.15/0.02 = 7.5$
Let $RR_{10} = p_{10}/p_{00} = 0.04/0.02 = 2$
Let $RR_{01} = p_{01}/p_{00} = 0.05/0.02 = 2.5$

We can also rewrite the multiplicative interaction for risk ratios as:

$RR_{11} / (RR_{10} \times RR_{01}) = [\ p_{11}/p_{10}\ ] / [\ p_{01}/p_{00}\ ]$
$= [\ 0.15/0.04\ ] / [0.05/0.02]$
$= 3.75 / 2.5$
$= 1.5$

Multiplicative Interactions

Data:

	E=0	E=1
G=0	0.02	0.05
G=1	0.04	0.15

In case control studies it is not in general possible to estimate risks or even risk ratios but one can still estimate odds ratios:

Let $OR_{11} = \{p_{11}/(1-p_{11})\} / \{p_{00}/(1-p_{00})\}$

Let $OR_{10} = \{p_{10}/(1-p_{10})\} / \{p_{00}/(1-p_{00})\}$

Let $OR_{01} = \{p_{01}/(1-p_{01})\} / \{p_{00}/(1-p_{00})\}$

Interaction on the odds ratio scale is then measured by:

$OR_{11} / (OR_{10} \times OR_{01})$

Again if this is > 1 the interaction is positive, if < 1 then negative

If the outcome is rare then the OR interaction measure approximates the RR interaction measure

Additive vs. Multiplicative Interactions

Conceived this way, interaction depends on the scale (multiplicative or additive)

We may in fact have additive interaction w/o multiplicative interaction

	E=0	E=1
G=0	0.02	0.05
G=1	0.04	0.10

Additive:

$p_{11} - p_{10} - p_{01} + p_{00} = 0.10 - 0.04 - 0.05 + 0.02 = 0.03 > 0$

Multiplicative:

$RR_{11} / (RR_{10} \times RR_{01}) = 5 / (2 \times 2.5) = 1$

Additive vs. Multiplicative Interactions

In other settings we may have multiplicative interaction but no additive interaction:

	E=0	E=1
G=0	0.02	0.05
G=1	0.07	0.10

Additive:

$p_{11} - p_{10} - p_{01} + p_{00} = 0.10 - 0.07 - 0.05 + 0.02 = 0$

Multiplicative:

$RR_{11} / (RR_{10} \times RR_{01}) = 5 / (3.5 \times 2.5) = 0.57 < 1$

Additive vs. Multiplicative Interactions

We can even have positive additive interaction but negative multiplicative interaction!

	E=0	E=1
G=0	0.02	0.05
G=1	0.07	0.12

Additive:

$p_{11} - p_{10} - p_{01} + p_{00} = 0.12 - 0.07 - 0.05 + 0.02 = 0.02 > 0$

Multiplicative:

$RR_{11} / (RR_{10} \times RR_{01}) = 6 / (3.5 \times 2.5) = 0.69 < 1$

Additive vs. Multiplicative Interactions

Thus we can have additive interaction without multiplicative
Or multiplicative interaction without additive
Or we can have both positive additive and multiplicative interaction
Or we can have both negative additive and multiplicative interaction
Or we can have positive additive and negative multiplicative!

Interaction is scale dependent

Moreover, there will almost always be interaction on some scale:
If both exposures have an effect we must have interaction on some scale either the additive or multiplicative scale (Rothman et al., 2008; VanderWeele, 2012)

Good practice to report on both scales; different scales are useful for different purposes
We should to identify motivations for studying interaction and choose the scale and type of interaction analysis appropriately
Even if we just want to examine differential effects, both scales should be used

Additive vs. Multiplicative Interactions

For some time in the epidemiologic literature, there was debate as to which scale one should assess interactions on (Blot and Day, 1979; Saracci, 1980; Rothman et al., 1980)

The general historical consensus was:

(1) Often the additive scale is of greatest public health importance
    It allows one to discern whether the effect would be different in different subgroups (Rothman et al., 1980)

(2) The additive scale also seemed to correspond to the more biological notion of synergism as conceived of by Rothman (1976) *[later in lecture]*

(3) However, sometimes the multiplicative scale (or neither scale) may be the one that more naturally corresponds to the biological mechanisms (Siemiatycki and Thomas, 1981); though in these cases the additive scale is still important for assessing public health impact

Both additive and multiplicative interaction measures should be reported

There is <u>no good reason</u> not to report additive as well

Statistical Interactions

Statistical interaction: A product term in a statistical model

Depending on the measure used, a statistical interaction could be additive or multiplicative, or it could be a component used to calculate an additive or multiplicative measure.

Statistical Interactions

A statistical model on the linear scale accommodating interaction takes the form:

$P(Y=1|G=g,E=e) = \alpha_0 + \alpha_1 g + \alpha_2 e + \alpha_3 eg$

In the regression setting:

$\alpha_3 = p_{11} - p_{10} - p_{01} + p_{00}$

i.e. the interaction contrast on the additive scale

In fact:

$\alpha_0 = p_{00}$

$\alpha_1 = p_{10} - p_{00}$

$\alpha_2 = p_{01} - p_{00}$

Statistical Interactions

Similarly one might have a "log-linear" model for risk ratios:
$\log \{P(Y=1|G=g,E=e)\} = \gamma_0 + \gamma_1 g + \gamma_2 e + \gamma_3 eg$

$\exp(\gamma_0) = p_{00}$ $\exp(\gamma_1) = RR_{10}$ $\exp(\gamma_2) = RR_{01}$
$\exp(\gamma_3) = RR_{11} / (RR_{10} \times RR_{01})$

Or one might alternatively use a "logistic" model for odds ratios:
$\text{logit} \{P(Y=1|G=g,E=e)\} = \gamma_0 + \gamma_1 g + \gamma_2 e + \gamma_3 eg$

$\exp(\gamma_0)=p_{00}/(1-p_{00})$ (cohort study only) $\exp(\gamma_1)=OR_{10}$ $\exp(\gamma_2) = OR_{01}$
$\exp(\gamma_3) = OR_{11} / (OR_{10} \times OR_{01})$

In these cases γ_3 is again referred to as a "statistical interaction" but now on the risk ratio or odds ratio scale (see *Appendix Slides* for "case-only" estimators of multiplicative interaction)

Statistical Interactions

The software used to fit models such as:

$$P(Y=1|G=g,E=e) = \alpha_0 + \alpha_1 g + \alpha_2 e + \alpha_3 eg$$
$$\text{logit}\{P(Y=1|G=g,E=e)\} = \gamma_0 + \gamma_1 g + \gamma_2 e + \gamma_3 eg$$

will give confidence intervals and p-values for the interaction coefficients, i.e. α_3 and γ_3 respectively.

These statistical models can also easily accommodate additional confounding variables or covariates in the model

However, the linear risk model often does not fit well with continuous covariates so logistic regression is typically used with covariates

As we will see below, we can still assess the direction of additive interaction if we fit logistic regression using alternative measures, e.g. RERI

Additive Interactions and RERI

Additive interaction $p_{11} - p_{10} - p_{01} + p_{00}$ divided by p_{00}
We get: $RR_{11} - RR_{10} - RR_{01} + 1$

This gives us something like the additive interaction but using RR's (or OR's); it is sometimes called the "Relative Excess Risk due to Interaction" or "RERI" (Rothman, 1986) or "Interaction contrast ratio" (Rothman, 2008)

If RERI > 0 we have a positive additive interaction
If RERI < 0 a negative additive interaction
We can thus assess additive interaction using risk ratios (or odds ratio if the outcome is rare)

If we fit a logistic regression we estimate OR's (approximate RR's for a rare outcome) and can obtain RERI:

With: $\text{logit}\{P(Y=1|G=g,E=e)\} = \gamma_0 + \gamma_1 g + \gamma_2 e + \gamma_3 eg$
Then: $RERI \approx OR_{11} - OR_{10} - OR_{01} + 1$
$= \exp(\gamma_1 + \gamma_2 + \gamma_3) - \exp(\gamma_1) - \exp(\gamma_2) + 1$

Additive Interaction and RERI

Consider again the data from Figueiredo et al. (2004) with odds ratios compared to the reference category G=0 (i.e. T/T or T/M) and E=0 (i.e. no alcohol):

	No Alcohol	Alcohol
T/T or T/M	1.00	1.12 (0.81-1.54)
M/M	1.21 (0.70-2.09)	2.09 (1.16-3.78)

Multiplicative Interaction: $OR_{11} / (OR_{10} \times OR_{01}) = 2.09 / (1.21 \times 1.12) = 1.54 > 1$

Also RERI $\approx OR_{11} - OR_{10} - OR_{01} + 1 = 2.09 - 1.21 - 1.12 + 1.00 = 0.76 > 0$
(under rare disease assumption)

There is a positive interaction on the additive scale

Other Additive Interactions

Another alternative measure for assessing additive interaction that is sometimes used is (Rothman, 1986):

Attributable Proportion (AP) of Interaction for Disease:

$$AP = (p_{11} - p_{10} - p_{01} + p_{00})/p_{11} = RERI / RR_{11}$$

i.e. how much of the *risk of disease* when both exposures are present is due to interaction (AP>0 positive additive interaction; AP<0 negative interaction)

Attributable Proportion (AP*) of Interaction for Effects:

$$AP^* = (p_{11} - p_{10} - p_{01} + p_{00})/(p_{11} - p_{00}) = RERI / (RR_{11} - 1)$$

i.e. how much of the *joint effect* of both exposures is due to interaction

If $p_{00} = p_{10} = p_{01} = 0.1$ and $p_{11} = 0.4$ then

$AP = 0.3 / 0.4 = 75\%$ and $AP^* = 0.3 / 0.3 = 100\%$

Both may be of interest

Attributable Proportion for Effects

In fact we can decompose the joint effect of both exposures as follow:

$p_{11} - p_{00} = (p_{10} - p_{00}) + (p_{01} - p_{00}) + (p_{11} - p_{10} - p_{01} + p_{00})$

i.e. we can decompose the joint effect of both exposures into a component that is due to: (i) just the first exposure alone, (ii) just the second alone, and that is due to (iii) their interaction

We can calculate what portion of the of the joint effect is attributable to:

(i) just the first exposure: $(p_{10} - p_{00})/(p_{11} - p_{00})$

(ii) just the second exposure: $(p_{01} - p_{00})/(p_{11} - p_{00})$

(iii) And their interaction: $(p_{11} - p_{10} - p_{01} + p_{00})/(p_{11} - p_{00})$

Note this is a decomposition of the *joint* effect of *both* exposures; we could also decompose the *total* effect of *one* exposure that is or is not attributable to interaction with the other *[Additional Slides]*

Attributable Proportion for Effects

If we divide these proportions by p_{00} we obtain:

$$\frac{RR_{10}-1}{RR_{11}-1}$$
$$\frac{RR_{01}-1}{RR_{11}-1}$$
$$\frac{RERI}{RR_{11}-1}$$

If we fit a logistic regression model to the data:

$$\text{logit}\{P(Y=1|G=g,E=e)\} = \gamma_0 + \gamma_1 g + \gamma_2 e + \gamma_3 eg + \gamma_4' c$$

we can obtain the three proportion measures as follows (if rare disease):

$$\frac{RR_{10}-1}{RR_{11}-1} \approx \frac{e^{\gamma_1}-1}{e^{\gamma_1+\gamma_2+\gamma_3}-1}$$
$$\frac{RR_{01}-1}{RR_{11}-1} \approx \frac{e^{\gamma_2}-1}{e^{\gamma_1+\gamma_2+\gamma_3}-1}$$
$$\frac{RERI}{RR_{11}-1} \approx \frac{(e^{\gamma_1+\gamma_2+\gamma_3}-e^{\gamma_1}-e^{\gamma_2}+1)}{e^{\gamma_1+\gamma_2+\gamma_3}-1}.$$

See VanderWeele and Tchetgen Tchetgen (2014) for SAS/Stata code and Mathur and VanderWeele (2018) for R code to do this automatically and for CI's

These methods that attribute proportions of effects to interaction may be a more intuitive way to report interaction on the additive scale

Attributable Proportion for Effects

Variants on chromosome 15q25 are associated with lung cancer (Hung et al., 2008) but their effects seem to vary by smoking status with possibly no effect among non-smokers (Li et al., 2010); here we dichotomize the exposures…

Using case-control data from MGH if we fit the logistic regression we get:
logit {P(Y=1|G=g,E=e)} = $\gamma_0 + \gamma_1 g + \gamma_2 e + \gamma_3 eg + \gamma_4' c$

$$\gamma_1 = 0.04 \text{ (95\% CI: } -0.33, 0.41), \gamma_2 = 1.33 \text{ (95\% CI: } 1.01, 1.64), \gamma_3 = 0.49 \text{ (95\% CI: } 0.08, 0.89)$$

Using the approach above we obtain:

$$\begin{aligned}
\frac{RR_{10}-1}{RR_{11}-1} &\approx 0.8\% \text{ (95\% CI: } -6.2\%, 7.7\%) \\
\frac{RR_{01}-1}{RR_{11}-1} &\approx 51.4\% \text{ (95\% CI: } 33.4\%, 69.4\%) \\
\frac{RERI}{RR_{11}-1} &\approx 47.8\% \text{ (95\% CI: } 33.3\%, 62.3\%).
\end{aligned}$$

Of the joint effects of the variants and smoking on lung cancer about 51% is due to smoking alone and 48% due to the interaction with the variants
These proportions are easier to interpret and move one away from “testing”

Interaction vs. Effect Heterogeneity

Recall that additive vs. multiplicative was a distinction concerning the numerical scale on which we measure interaction strength in terms of association

A second important distinction, causal interaction vs. effect heterogeneity, concerns what kind of causal conclusions, if any, we can draw from numerical measures of interaction.

Interaction vs. Effect Heterogeneity

Two Sets of Confounders
With assessing interaction we have two exposures
We thus have 2 potential sets of confounders (for G and for E)
We might control for confounding for both, or just G, or just E, or neither

If we control for one set of confounders we can speak of the causal effects of one exposure; if we have controlled for both sets of confounders we can speak of causal effects of both

Effect Heterogeneity: The effect of one exposure differs across strata of another (we need to control for one set of confounders)
Causal Interaction: The effect of one exposure would differ if we intervened to fix the second exposure to a different value (we need to control for both sets of confounders)

Interaction vs. Effect Heterogeneity

Counterfactual Notation:

Let Y_e denote the counterfactual outcome if we set E to e

Let Y_{eg} denote the counterfactual outcome if we set E to e and G to g

When thinking about interpreting effect estimates from interaction analyses we need to think carefully about whether we have controlled for confounding for only one or both relationships

Do our estimates refer to causal effects of one or both exposures?

e.g. In a gene-environment interaction study, have we controlled for confounding for both the environmental factor and the genetic factor?

Interaction vs. Effect Heterogeneity

In general researchers simply run a regression of the outcome Y on the exposure E, G and a product term E*G, possibly also controlling for some covariates C

However, whether the regression coefficient for the product term E*G can be interpreted as a measure of effect heterogeneity or causal interaction or both or neither depends on the causal relationships governing the variables E, G, Y and C.

For effect heterogeneity we only have one set of confounding factors to consider, just for the E-Y relationship

For causal interaction we have two sets of confounding factors to consider, those for the E-Y and for the G-Y relationships

Sometimes, measures of causal interaction and effect heterogeneity will coincide (e.g. when C blocks all backdoor paths from G to Y) but at other times they will not.

Example: Homelessness

Intervention: Supportive Housing for Homeless (E)

Secondary Variable: Subgroup Analyses by Part-Time Employment (G)

Outcome: Hospitalizations (Y)

Observational Study:

Perhaps the data come from an observational study and homeless individual with more severe chronic conditions (C=1) are more likely to receive the supportive housing (E)

Results:

Maybe we find an overall effect of E on Y

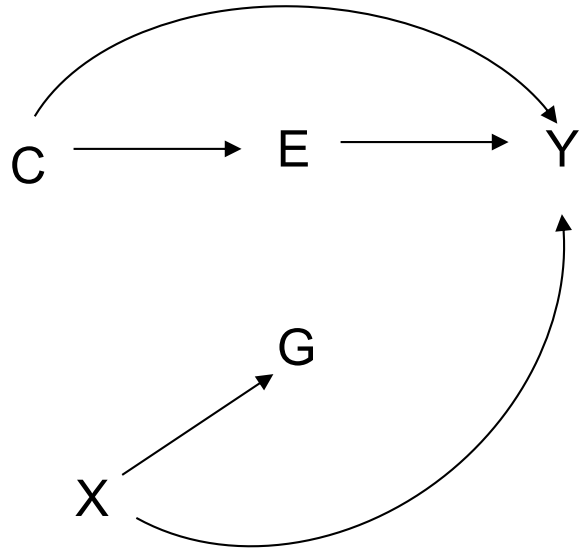
Maybe we then find the effect of Y is larger when G=1

We can *target* those with part-time employment (G=1) for a larger effect

Should we also *intervene* to provide part-time employment to enhance the the effect of the supportive housing?

Interaction vs. Effect Heterogeneity

Y – Hospitalization
E – Housing Intervention
C – Chronic Conditions
X – Mental Health
G – Employment



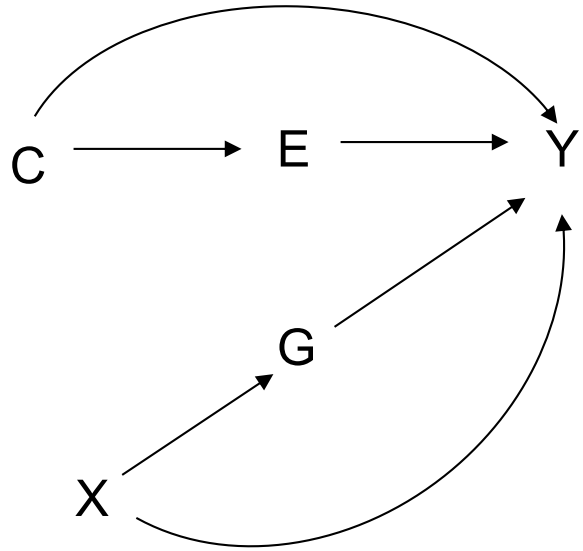
<u>Suppose employment were a proxy for mental health status</u> which may be what truly interacts with the housing intervention

In our study, we found a larger effect for the subgroup with employment because the group with employment had better mental health status

However, intervening on employment while leaving mental health unchanged will not change the effect of treatment

Interaction vs. Effect Heterogeneity

Y – Hospitalization
E – Housing Intervention
C – Chronic Conditions
X – Mental Health
G – Employment



Even if there were an effect of employment on hospitalizations, in order to estimate this effect of employment, we would have to control for X

If we control only for C, we have obtain estimates of effect heterogeneity for E across strata of G

If we control for (C,X) we can obtain estimates of causal interaction (i.e. what would happen if we intervened on E and on G)

Effect Heterogeneity

We may be interested in how the treatment effects vary by strata of G:

$E[Y_1 - Y_0 \mid G = g]$ (”effect heterogeneity”/“effect modification” if this varies with g)

This would be of interest in **choosing groups to target** for treatment

If both E and G are binary, then as a measure of effect heterogeneity on the risk difference scale we would use: $E[Y_1 - Y_0 \mid G = 1] - E[Y_1 - Y_0 \mid G = 0]$

If the effect of E on Y is unconfounded conditional on G we have:

$E[Y_1 - Y_0|G=1] - E[Y_1 - Y_0|G=0] = p_{11} - p_{10} - p_{01} + p_{00}$

We can define measures of effect heterogeneity on multiplicative scales also:

e.g. $\{P(Y_1|G=1)/P(Y_0|G=1)\} / \{P(Y_1|G=0)/P(Y_0|G=0)\}$

We could also obtain similar results conditional on measured covariates C

Causal Interaction

We may instead be interested in whether the treatment effect would vary if we were to intervene on G

$E[Y_{1g} - Y_{0g}]$ ("causal interaction" if this varies with g)

We might be interested in causal interaction if we were thought we might **intervene** on G to make treatment E more effective or if we were interested in the mechanisms by which E and G interacted

If both E and G are binary then as a measure of causal interaction we would use: $E[Y_{11} - Y_{01}] - E[Y_{10} - Y_{00}]$

If the effects of E and G on Y are unconfounded we have:

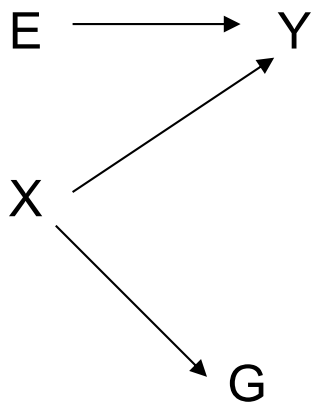
$E[Y_{11} - Y_{01}] - E[Y_{10} - Y_{00}] = p_{11} - p_{10} - p_{01} + p_{00}$

We can define measures of causal interaction on multiplicative scales also:

e.g. $\{P(Y_{11})/P(Y_{01})\} / \{P(Y_{10})/P(Y_{00})\}$

Effect Heterogeneity w/o Interaction

Y – Hypertension
E – Drug for hypertension
X – Genotype
G – Hair color



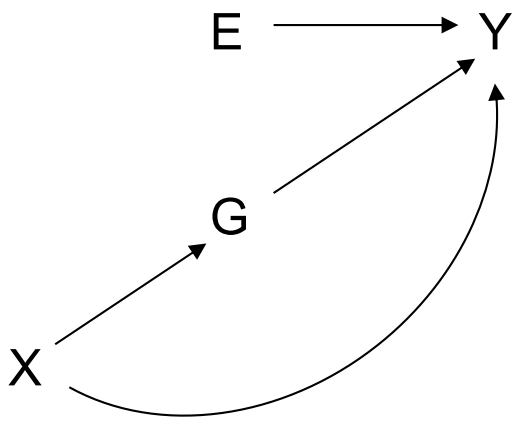
Suppose G has no effect on Y so there is no causal interaction between E and G on Y

But G may still be an effect modifier for Y because conditioning on G gives information on X which does affect Y

We thus see that we can have effect heterogeneity with respect to G without causal interaction with respect to G

Interaction w/o Effect Heterogeneity

Y – Weight of obese children
E – Drug for weight loss
G – Exercise
X – Sugar intake



Suppose the effect of the drug E and exercise G might interact to affect the weight of obese children; but sugar intake X suppresses the effect

If we were to simply look at the effect of E within strata of G, high levels of exercise might also indicate high levels of sugar intake which increases both exercise G and weight Y

These effects of sugar X on weight Y (through exercise and directly) might cancel so we observe constant effect of E within strata of G

Therefore we can have <u>interaction with no effect heterogeneity</u>

Subgroup Analysis in RCTs

RCT Interpretation: In a randomized trial of treatment E when we look at subpopulations we can always interpret results as effect heterogeneity (the intervention A is randomized so there is no confounding for E)

Confounding and Subgroups: However, we will not have in general randomized the secondary factor so confounding may be an issue

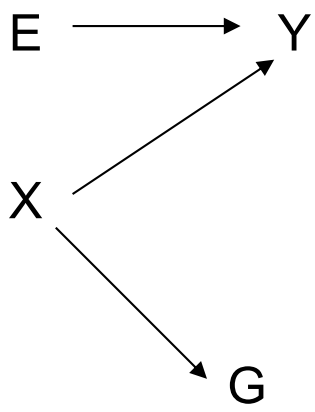
We can only interpret estimates as those of "causal interaction" under additional unconfoundedness assumptions, when the confounders for the secondary factor G are also controlled for

If we are simply interested in targeting populations in which the effect of the intervention is as large as possible then effect modification will generally be what is of interest (we are only considering one intervention across strata of another variable)

If we are interested in assessing possible interventions on two exposures then we want measures of causal interaction

Subgroup Analysis in RCTs

Y – Hospitalization
E – Housing Intervention
X – Mental Health
G – Employment



Suppose the supportive housing intervention E were randomized

Suppose once again we found larger effects of those with employment

Once again, if employment were a proxy for mental health status then intervening on employment while leaving mental health unchanged will not change the effect of treatment

Subgroup Analysis in RCTs

Knowing that the effect is larger in the subgroup with employment could help us target populations in which the housing intervention would have a larger effect

But if we are going to intervene on the factor that defines the subgroups we need to know that this factor itself (and not simply something associated with it) is what alters the effect of the intervention

Again we need to distinguish between mere effect heterogeneity and causal interaction…

A Note on Terminology

In practice, epidemiologists use "effect modification" for both effect heterogeneity and causal interaction; the terms "interaction" and "effect modification" are themselves often used interchangeably

In the social sciences "moderation" is often used for both

There is some ambiguity in terminology; however, even if the terms themselves are used interchangeably, there are still two distinct concepts present and researchers need to be aware of this

The distinction again has to do with whether one or two potential interventions are in view

Failure to take the distinction into account could lead to incorrect policy recommendations

Without committing to terminology the distinctions can be acknowledged in the methods and discussion sections of a paper

Presenting Interaction

We have discussed various methods to go about estimating and drawing inference for additive interaction

We will briefly consider how best to present interaction analyses

Suppose that we were going to report effect estimates on a risk ratio scale and that we had the following risks:

	E=0	E=1
G=0	0.02	0.05
G=1	0.07	0.10

The Problem of Separate Reference Groups

One approach to presenting interaction analyses would be to use a separate reference group for the G=0 and G=1 strata

	E=0	E=1
G=0	RR=1 (ref)	RR=2.5
G=1	RR=1 (ref)	RR=1.43

Using this approach we can see that the risk ratio for E is different in different strata of G; we can even calculate multiplicative interaction: 1.43/2.5= 0.57

But… do we know whether the (G=0,E=1) or (G=1,E=1) subgroup has higher risk?
Do we know whether the (G=0,E=0) or (G=1,E=0) subgroup has higher risk?
Do we know whether the (G=0,E=1) or (G=1,E=0) subgroup has higher risk?
Can we estimate additive interaction?

The Problem of Separate Reference Groups

None of this can be determined from the relative risks when separate reference groups are used for each stratum

Suppose instead we used a common reference group (e.g. G=0,E=0). We then have:

	E=0	E=1
G=0	RR=1 (ref)	RR=2.5
G=1	RR=3.5	RR=5.0

We can still calculate the risk ratios for E by G (e.g. RR=2.5 and RR=5/3.5=1.43)
We can estimate multiplicative interaction: 5.0/(2.5 x 3.5) = 0.57
We can estimate additive interaction 5.0 - 3.5 - 2.5 + 1 = 0
We know that the ordering of the categories by risk is (G=0,E=0), (G=0,E=1), (G=1,E=0), (G=1,E=1)

Presenting interaction analyses with a common reference group is thus much more information for the reader
If we think it may be helpful we can also report in the table RR's for E by G (or for G by E), along with measures of additive and multiplicative interaction

The Problem of Separate Reference Groups

Consider analyses of religious service attendance and Protestant vs. Catholic denomination with either suicide (VanderWeele et al., 2016) or divorce (Li et al., 2018) as outcomes.

Suppose we use separate reference groups… We then have:

	Suicide		Divorce/Separation	
	Protestant	Catholic	Protestant	Catholic
<1/week	RR=1 (ref)	RR=1 (ref)	RR=1 (ref)	RR=1 (ref)
>=1/week	RR=0.34	RR=0.05	RR=0.62	RR=0.48

It looks like the “effect” of attendance is larger for Catholics, but with a single reference:

	Suicide		Divorce/Separation	
	Protestant	Catholic	Protestant	Catholic
<1/week	RR=1 (ref)	RR=0.97	RR=1 (ref)	RR=1.29
>=1/week	RR=0.34	RR=0.05	RR=0.62	RR=0.65

The explanations for the ”larger effect” for Catholics look quite different: more protective for suicide, but for divorce, higher baseline rate for Catholics

Presenting Interaction

The STROBE guidelines for observational studies (von Elm et al., 2007) thus recommend that when interaction analyses are presented a common reference group is used

In general it is good to report all of the following (Knol and VanderWeele, 2012):
(i) Stratum-specific risk ratios and CI's with a common reference group
(ii) Risk ratios and CI's for E by strata of G (and possibly for G by strata of E)
(iii) Measures and CI's for multiplicative interaction and additive interaction
(iv) The confounding variables for which control has been made (and whether these are intended to control for confounding of one or both factors)

This may seem reasonably straightforward but it is often not done
In the random sample of studies examined by Knol et al. (2009), of those that did report interaction, only about 10% used a common reference group (about 70% presented stratum specific effects, and 20% only statistical significance)
Almost none of the studies reported on additive interaction

Now you know how to do so…

General Conclusions

(1) Interaction is scale dependent

(2) If both exposures have an effect, there will always be interaction on some scale

(3) It is best practice to assess interaction on both additive and multiplicative scales

(4) It is best practice to report interaction with a single reference group

(5) We need to think carefully about confounding in interaction analyses

Part II. Motivations for Studying Interaction and Forms of Interaction

Motivations for Interaction Analysis

(1) Selecting subgroups to treat when resources are limited

(2) Identifying subgroups in which treatment may be harmful

(3) Understanding mechanisms for the outcome

(4) Assessing the generalizability of an effect estimate

(5) When it is not possible to intervene on the exposure, interactions can help identify effect modifiers to intervene upon to eliminate exposure effects

(6) Assessing whether the effect is "bigger" for one subgroup vs. another *[but, as in last lecture, the very question can be scale dependent]*

[Note: We might include a product term in a statistical model not because we are interested in "interaction" but simply to add model flexibility]

Motivation 1. Selecting Subgroups

If resources are limited we may only be able to intervene on one of two subgroups

Interactions could help determine which of these two subgroups to choose in order to the most public health impact

Note: Here we are thinking of interventions on one exposure across strata defined by another variable (i.e. effect heterogeneity)
With observational data we need the effect of our primary exposures on the outcome to be unconfounded

To evaluate for which of two subgroups an intervention would maximize public health impact, we can assess this with additive interaction

To see why, let us consider an example…

Public Health Importance

Suppose that E denotes some drug and the outcome is "survival at 3 years":
And there are 100 with G=0 and 100 with G=1 and we have 100 doses

	E=0	E=1
G=0	0.01	0.05
G=1	0.04	0.10

The risk difference for E on those with G=0 is: 0.05 - 0.01 = 0.04
The risk difference for E on those with G=1 is: 0.10 - 0.04 = 0.06
$p_{11} - p_{10} - p_{01} + p_{00}$ = 0.10 - 0.04 - 0.05 + 0.01 = 0.02 > 0

The risk ratio for E on those with G=0 is: 0.05 / 0.01 = 5
The risk ratio for E on those with G=1 is: 0.10 / 0.04 = 2.5
$RR_{11} / (RR_{10} \times RR_{01})$ = 10 / (5 x 4) = 0.5 < 1

If we give the drug to G=0 group the number surviving is: 100*(0.05) + 100*(0.04) = 9
If we give the drug to G=1 group the number surviving is: 100*(0.01) + 100*(0.10) = 11

We should treat the G=1 group; we have an additional 2 persons surviving
Additive interaction (not multiplicative interaction) identifies this, and always does so
The risk ratio suggests treating the G=0 group; for public health purposes we should rely on additive interaction measure to decide which subgroups to target

Additive vs. Multiplicative Interactions

In most published epidemiologic studies, interactions are evaluated and reported only on the multiplicative scale

Interaction on the additive scale are often not e.g. perhaps only about 1 in 50 in epidemiology reported (Knol et al., 2009), even though there has been consensus that it should be reported for assessing public health relevance

The focus on multiplicative interaction is likely due to the statistical models which are used in such analyses (e.g. logistic regression) and the fact that the models employed immediately give interactions (and confidence intervals) on a multiplicative scale

In general, if interaction is of interest it is good to report estimates and confidence intervals on both scales

The economics literature has been more consistent about assessing interaction on the additive scale to assess public health impact

Additive Interaction with Logistic Regression

For a rare outcome, recall that we can still obtain measures of additive interaction if we fit logistic regression using RERI:

$$\text{logit}\{P(Y=1|G=g,E=e)\} = \gamma_0 + \gamma_1 g + \gamma_2 e + \gamma_3 eg$$

We can then calculate:

$$\text{RERI} \approx OR_{11} - OR_{10} - OR_{01} + 1$$
$$= \exp(\gamma_1 + \gamma_2 + \gamma_3) - \exp(\gamma_1) - \exp(\gamma_2) + 1$$

If RERI > 0 treatment in the G=1 group maximizes impact
If RERI < 0 treatment in the G=0 group maximizes impact

Logistic Regression and RERI

For RERI, Hosmer and Lemeshow (1992) give standard errors using the delta method

Lundberg et al. (1996) provides some SAS code but requires recoding of the exposures

Norton et al. (2004) give Stata code for estimating additive measures of interaction from logit and probit models

VanderWeele and Knol (2014) provide easy-to-use SAS and Stata code to implement this directly without using a spreadsheet or recoding exposures; Mathur and VanderWeele (2018) provide R code

Statistical Interactions

For continuous exposures G and E we can simply replace exposure levels 0 and 1 by g_0 and g_1, and/or e_0 and e_1 and proceed as before

For RERI we could fit the logistic model:
logit {P(Y=1|G=g,E=e)} = $\gamma_0 + \gamma_1 g + \gamma_2 e + \gamma_3 eg$

We would then use as the estimate of RERI:

$$e^{(g_1-g_0)\gamma_1+(e_1-e_0)\gamma_2+(g_1e_1-g_0e_0)\gamma_3} - e^{(g_1-g_0)\gamma_1+(g_1-g_0)e_0\gamma_3} - e^{(e_1-e_0)\gamma_2+(e_1-e_0)g_0\gamma_3} + 1.$$

SAS and Stata code is given in VanderWeele and Knol (2014); standard errors can be obtained with the delta method

Note: Different changes in the exposure may have different measures of interaction; see Knol et al. (2007) for further discussion

The same approach can be used with hazard ratios with a rare outcome by replacing the odds ratio with hazard ratios (Li and Chambless, 2007)

Motivation 2. Avoiding Harm in Subgroups

In some settings an exposure may have a positive effect on one subgroup but a negative effect on a different subgroup

Even if resources were not limited we would not want to treat everyone
For some individuals the treatment would be harmful

Such cases are referred to as “qualitative” or “crossover” interaction

It is very important to detect such interaction if present

These qualitative or crossover interaction are independent of scale
If they are present on the additive scale they will be on the multiplicative as well and vice versa

A special care of “qualitative interaction” is when there is an effect in one subgroup but no effect at all in another; this is sometimes called a “pure” interaction; such pure interactions are also independent of scale

Qualitative Interaction

Gail and Simon (1985) consider data from a trial of two therapies for breast cancer, one of which does and the other of which does not involve tamoxifen

For young patients under age fifty with low progesterone receptor levels, the treatment without tamoxifen led to higher proportions who were disease-free at three years

For all other groups (who were either older, or had higher progesterone receptor levels, or both) the treatment with tamoxifen led to higher proportions who were disease-free at three years

Here we would likely want to give young patients with low progesterone receptor levels the treatment without tamoxifen, and others the treatment with tamoxifen

Note: here we only need the effect of the primary exposure unconfounded

Qualitative Interaction

Several statistical approaches have been developed for testing for such qualitative interaction

The details of the approaches and power differ but they all basically coincide when there are only two subgroups

The approaches with three or more subgroups become more complicated

When testing for qualitative interaction across two subgroups one particularly simple approach (Pan and Wolfe, 1997) to test for a qualitative interaction at the 5% significance level is to construct 90% confidence intervals for the exposure effect in each of the two subgroups.

If, on a difference scale say, one of the 90% confidence intervals lies entirely above 0 and the other lies entirely below 0, then one would reject the null hypothesis of no qualitative interaction

Note that only 90% confidence intervals (not 95%) need to be constructed

Qualitative Interaction

Pan and Wolfe (1997) describe an approach for multiple subgroups and allows also testing for qualitative interaction of at least a certain magnitude (rather than simply larger versus small than zero)

It essentially just requires constructing confidence intervals of various sizes depending on the number of subgroups.

Their approach is equivalent to that described by Piantadosi and Gail (1993), sometimes referred to as the "range test," but easier to implement

An alternative approach was proposed by Gail and Simon (1985) which involves not simply constructing confidence intervals for the effects in each subgroup but rather constructing a confidence interval for the sum of the positive versus negative standardized effects across subgroups.

The Gail and Simon (1985) approach tends to perform better when there are several subgroups with positive effects and several with negative

The Piantadosi and Gail (1993) and Pan and Wolfe (1997) approach tends to perform better if the effects in most of the subgroups are in one direction and there are only one or very few subgroups with opposite effects

Optimal Subgroup Selection

How does this generalize to a covariate vector C?

We want to divide population into treated (T) and untreated (S) to maximize outcomes: $E(Y_1|T)P(T) + E(Y_0|S)P(S)$

The optimal treatment rule with covariates C=c is:

$$T = \{c : E[Y|A=1,c] - E[Y|A=0,c] > 0\}$$

How large this is will vary with how rich C is and how well we can model E[Y|a,c]; different machine learning approaches can be compared

Statistical inference in this setting is challenging using the same data to model E[Y|a,c] to determine who to treat, and to estimate outcomes under the treatment rule (cf. Cai et al., 2011; Zhao et al., 2013; Luedke and van der Laan, 2016; VanderWeele et al., 2019)

We essentially simultaneously have the difficulties of detecting interaction (generally low power) and something analogous to multiple testing with C

Note: this is different from stratifying on "prognostic scores": E[Y|A=0,c]

Motivation 3. Mechanistic Interaction

Do statistical interactions tell us anything about biological or mechanistic interactions?

Several authors have pointed out the potential danger of using statistical interaction to draw conclusions about biological interaction (Siemiatycki and Thomas, 1981; Thomas, 1991; Rothman and Greenland, 1998; Cordell, 2002)

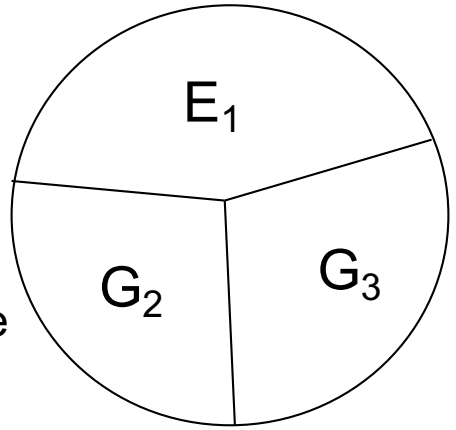
How might we conceive of mechanistic interaction?

Mechanistic Interaction

Sufficient Causation in Statistics and Epidemiology

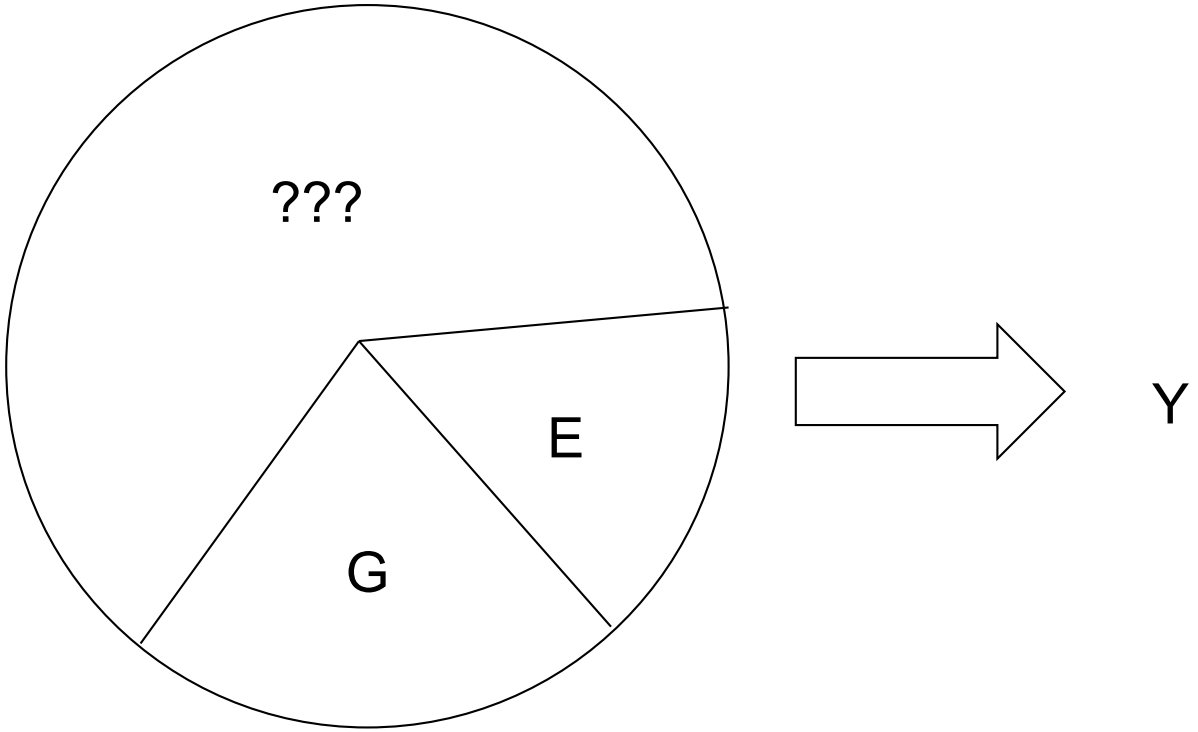
Rothman (1976) defined a "sufficient cause" as minimal set of events, conditions or characteristics that inevitably produce the disease; a "component cause" (or "cause") was an individual event, condition or characteristic required by a given sufficient cause.

- Rothman also provided a schematic for these component causes which have come to be known as "causal pies"
- There has been some work relating sufficient component causes to potential outcomes (Greenland and Poole 1988, Rothman and Greenland 1998, Aickin 2002, Flanders 2006, VanderWeele and Hernan 2006, VanderWeele and Robins 2007, 2008)
- Similar ideas also appeared earlier in the philosophical literature (Mackie, 1965)



Mechanistic Interaction

We may want to know whether two causes G and E are ever both present in the same sufficient cause



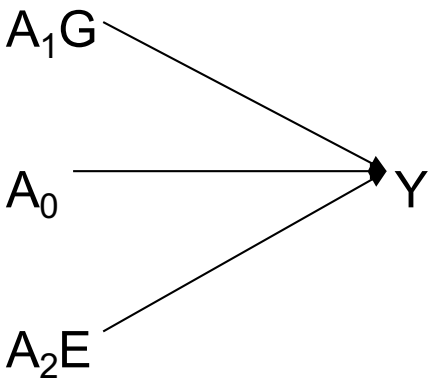
Can we learn anything from statistical interaction...?

Negative multiplicative interaction with no mechanistic interaction

Suppose there are just three mechanisms for outcome Y

- One requiring G and some other factors A_1
- One requiring E and some other factors A_2
- One requiring neither G nor E, just some other factors A_0

Suppose that G and E are independent in the population



There is no mechanism requiring both G and E

There is no interaction between G and E in a mechanistic sense

Negative multiplicative interaction with no mechanistic interaction

Suppose G, E, A_1, A_2 and A_0 are all independent in the population

Suppose G and E occur with probabilities 0.2 and 0.5 respectively

Suppose A_1 and A_2 each occur with probability 0.015

Suppose A_0 occurs with probability 0.005

Suppose there are 10,000 subjects in the population

G	E	Total	Cases	RR
G=0	E=0	4000	20	1
G=1	E=0	1000	20	4.0
G=0	E=1	4000	80	4.0
G=1	E=1	1000	35	7.0

Multiplicative interaction:

$RR_{11} / (RR_{10} \times RR_{01}) = 7.0/(4 \times 4)$

$= 0.44 < 1$

RERI = 7.0 - 4.0 - 4.0 + 1 = 0

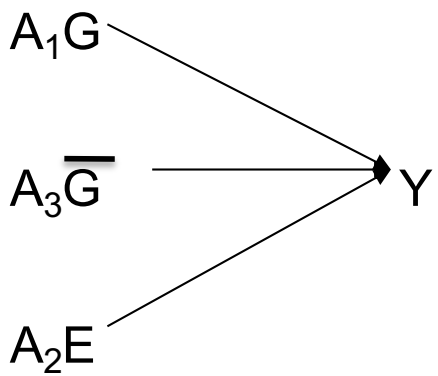
We have a non-zero negative multiplicative interaction but… This can arise without any sort of mechanistic interaction! i.e. no mechanism with both G and E

Positive multiplicative interaction with no mechanistic interaction

Suppose there are three mechanisms for outcome Y

- One requiring G and some other factors A_1
- One requiring E and some other factors A_2
- One requiring *absence* of G (denoted by $\overline{G}$) and factors A_3

Suppose that G and E are independent in the population



There is no mechanism requiring both G and E

There is no interaction between G and E in a mechanistic sense

Positive multiplicative interaction with no mechanistic interaction

Suppose G and E are independent and occur with probabilities 0.4 and 0.5 respectively
Suppose distribution of A_1, A_2 and A_3 is as follows

$P(A_1 =0, A_2 =1, A_3 =1) = .004$
$P(A_1 =1, A_2 =0, A_3 =0) = .004$
$P(A_1 =0, A_2 =1, A_3 =0) = .001$
$P(A_1 =1, A_2 =1, A_3 =0) = .001$
$P(A_1 =0, A_2 =0, A_3 =0) = .99$

Suppose there were 10,000 subjects in the population

G	E	Total	Cases	RR
G=0	E=0	3000	12	1
G=1	E=0	2000	10	1.25
G=0	E=1	3000	18	1.50
G=1	E=1	2000	20	2.50

$RR_{11} / (RR_{10} \times RR_{01})$

$= 2.50 / (1.25 \times 1.50) = 1.33 > 1$

$RERI = 2.50 - 1.50 - 1.25 + 1 = 0.75 > 0$

We have positive multiplicative and additive interaction

But there is no mechanistic interaction

Notation and Definitions

When *can* we conclude that there is a mechanism that requires both exposures?

Counterfactuals: Let Y_{ge} denote the counterfactual outcome for an individual if, possibly contrary to fact, G is set to g and E is set to e

Causal Interdependence: We say that there is causal interdependence if for some individual $Y_{11}=1$ but $Y_{10}=Y_{01}=0$

It can be shown that if such "causal interdependence" is present then there must be "synergism" in Rothman's sufficient cause framework (VanderWeele and Robins, 2008) i.e. a sufficient cause with both G and E

Example: Are there individuals who would develop breast cancer with the XRCC3-T241M risk allele and alcohol consumption but not if only one or the other were present?

Mechanistic Interaction

Such sufficient cause interaction is not equivalent to statistical interaction (Greenland and Poole, 1988; Rothman and Greenland, 1998)

Testing for such sufficient cause interaction in general requires stronger assumptions than statistical interaction

Monotonicity: We will then say that G has a positive "monotonic effect" on the outcome Y if Y_{ge} is non-decreasing in g (similarly for E)

Monotonicity requires the effect always operates in the same direction for all individuals; it might be plausible sometimes (e.g. the effect of smoking on lung cancer) but not others (e.g. alcohol on stroke)

Mechanistic Interaction

Let $p_{ge} = P(Y=1|G=g,E=e)$

Rothman and Greenland (1998) show that if the effects of G and E on Y are unconfounded and if both G and E have positive monotonic effects on the outcome then one can test for a sufficient cause interaction by testing:

$$p_{11} - p_{10} - p_{01} + p_{00} > 0$$

i.e. positive additive interaction [the effects of both exposures combined exceed the sum of the effects of each considered separately]

We could also test this by RERI > 0

Mechanistic Interaction

Result (VanderWeele and Robins, 2007, 2008): If the effects of G and E on Y are unconfounded one can test for a sufficient cause interaction by testing:

$$p_{11} - p_{10} - p_{01} > 0$$

This condition can be expressed as RERI > 1

It is a stronger condition than simply having positive additive interaction which would only require RERI > 0

By using this stronger condition one can in fact test for sufficient cause interaction even without monotonicity, contrary to what was previously thought (Rothman and Greenland, 1998)

Application 1: Breast Cancer

Figueiredo et al. (2004) studied the effects of XRCC3-T241M polymorphisms and various environmental factors on breast cancer risk

	No Alcohol	Alcohol
T/T or T/M	1.00	1.12 (0.81-1.54)
M/M	1.21 (0.70-2.09)	2.09 (1.16-3.78)

Here RERI $\approx$ OR_{11c} - OR_{10c} - OR_{01c} + 1

= 2.09 - 1.21 - 1.12 + 1.00 = 0.76 > 0

The estimate suggests evidence for sufficient cause interaction with the assumption that both alcohol and the M/M polymorphism have monotonic effects on the outcome since RERI>0 but that we cannot draw conclusions without monotonicity since RERI<1

Here, however, it is not very clear whether monotonicity will hold
Moreover, without access to the data we cannot calculate standard errors here

Application 2: Diarrheal Disease

Data from a case-control study in Northwestern Ecuador (2003-2008) indicates that (Bhavnani et al., 2012):

Giardia
Rotavirus
E. coli/Shigella

are all associated with increased risk of diarrheal disease

There is also strong evidence of interaction:

Giardia & rotavirus: RERI=10.7-2.6-1.1+1 = 7.9 (95% CI: 3.1, 18.9)
E. coli/Shigella & rotavirus: RERI=13.2-2.6-1.6+1 = 9.9 (95% CI: 2.6, 28.4)
E. coli/Shigella & giardia: RERI= 3.0-1.1-1.6+1 = 1.2 (95% CI: -1.4, 3.1)

For Giardia & rotavirus and for E. coli/Shigella & rotavirus, there is strong evidence of mechanistic interaction even without making any monotoncity assumptions (VanderWeele, 2012)

Some Caveats

Note: These tests are sufficient but not necessary; if the RERI condition fails there may still be sufficient cause interactions or there may not be; one cannot determine this from the data; the tests here, however, are as weak as possible without further assumptions

Note: If the conditions are fulfilled, then sufficient cause is present; but this means that there are some individuals for whom there is a sufficient cause or epistatic response pattern, not that this holds for all individuals

Note: If we use the contrasts in terms of absolute risks for sufficient cause interaction, then these provide *lower bounds* on the prevalence of individuals with such mechanistic interaction

Note: Using event stronger conditions we can also test for individuals with response patterns of individual $Y_{11}=1$ but $Y_{10}=Y_{01}=Y_{00}=0$ (Appendix Slides)

Limits of Biologic Inference

Suppose that G_1 and G_2 are two genetic factors
Suppose that when G_1=1 protein 1 is not produced
Suppose that when G_2=1 protein 2 is not produced

Suppose that the outcome Y occurs if and only if neither protein 1 nor protein 2 are present

We then have a sufficient cause interaction: the outcome occurs if and only if G_1=1 and G_2=1

But we do not have physical interaction here
It is precisely the absence of the proteins that gives rise to the outcome (there is nothing to physically interact here)

It is important to understand the limits of the conclusions being drawn about these alternative forms of causal interaction

Terminology

Sufficient cause interaction was sometimes earlier referred to as "biologic interaction" (e.g. Rothman and Greenland, 1998); and sometimes just additive interaction was even referred to as "biologic interaction" (Andersson et al., 2005)

As we have seen, neither statistical interaction nor even sufficient cause interaction necessarily tells us anything about physical interactions

Statistical analyses can only tell us limited information about the underlying biology (Siemiatycki and Thomas, 1981; Thomas, 1991; Rothman and Greenland, 1998; Cordell, 2002)

Because of this there has been a suggestion to move away from the use "biologic interaction" for sufficient cause interactions (cf. Lawlor, 2011; VanderWeele, 2011)

It may be more appropriate to refer to these sufficient cause or "epistatic interactions" (in genetics) as "mechanistic interactions" (both exposures together turns the outcome 'on' and the removal of one turns the outcome 'off')

Motivation 4. Generalizability

Some have argued that exposure risk ratios are more constant across other factors and across populations than risk differences

This is ultimately a claim that underlying human pathophysiology and biology results in multiplicative models with no interaction

If it is the case that exposure risk ratios are more constant across other factors and across populations than risk differences then this has important implications for generalizability

If the risk ratio is constant across all other risk factors (multiplicative model with no interaction) then we might try to use the same risk ratio to predict the effects of interventions in populations with very different risk factor distributions

Stability of Risk Ratios?

Evidence that is often presented comes from syntheses of meta-analyses suggesting the tests reject “risk difference heterogeneity” more often than “risk ratio heterogeneity” (Engels et al., 2000; Sterne and Egger, 2001; Deeks, 2002)

But this evidence ignores that these tests have very different power (Poole et al. 2015)

Anecdotal evidence also suggests that with environmental and genetic exposures, very often there is little evidence for multiplicative interaction (Spiegleman and VanderWeele, 2017)

More convincing evidence comes from careful modeling of many different genetic variants and environmental factors which suggest multiplicative models with no interaction often fit very well (Maas et al., 2016)

Multiplicative Interaction

Maybe the claim that human pathophysiology and biology is such that it often results in multiplicative models is true
More rigorous evidence is needed, but if it is so…

Then substantial multiplicative interaction (perhaps especially a substantial positive multiplicative interaction) may indicate that something is unusual about the underlying mechanisms

It is still a leap to draw conclusions underlying biological mechanisms, but it is one argument for examining multiplicative interaction

If there is no multiplicative interaction (the risk ratio is constant across all other risk factors) then we might try to generalize to other populations

But this still relies on there not being any "unmeasured effect modifiers" (e.g. a different city with a toxic exposure not present in the original study population which modifies the effect of the primary exposure)

Generalizability

The claim of near-uniform multiplicativity itself needs further evidence

But if RR's are constant this is an intriguing approach to extrapolating results… if it is true….
Many clinicians simply assume it holds

We always need to be cautious but, if it does hold, implications for generalizability are profound

Methods are also available when interactions are present to extrapolate effect estimates other populations with very different risk factor distributions (e.g. Cole and Stuart, 2010; Dahabreh et al., 2020)

Methods still require no unmeasured effect modifiers, though sensitivity analysis for this has now be developed (Nguyen et al., 2017)

Motivation 5. Attributing Effects to Interactions

When it is not possible to intervene on the exposure, interactions can help identify effect modifiers to intervene upon to eliminate exposure effects

See *Appendix Slides* for methods and measures

Motivation 6. Effect Magnitudes

Whether the effect is “bigger” in one subgroup versus another is scale-dependent:

- We can have additive interaction, but no multiplicative
- Or multiplicative interaction, but no additive
- Or we can have positive additive and negative multiplicative!

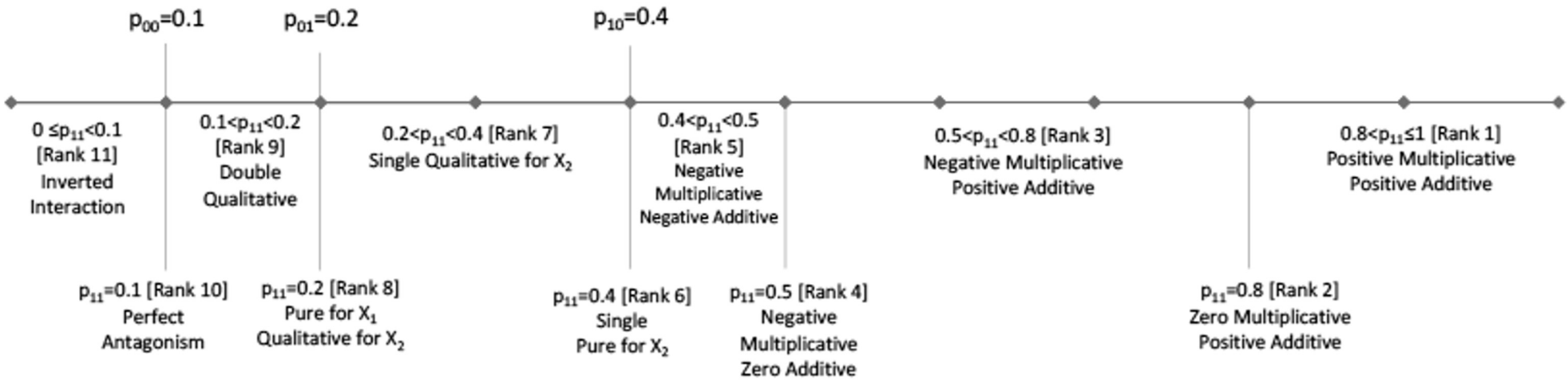
We only really can unambiguously state that the ”effect is larger” for on subgroup versus another if this is true e.g. on both scales (e.g. additive and multiplicative interaction are both positive, or additive and multiplicative interaction are both negative)

Interaction Continuum

More generally, in settings in which both exposures alone are in the causative direction (p_{10}-$p_{00}\geq 0$, p_{01}-$p_{00}\geq 0$), we can rank order different types of additive and multiplicative interaction patterns considered jointly

With p_{00}, p_{01}, p_{10} considered fixed, then as p_{11} descends, the interaction contrast moves through the ranks of an interaction continuum (VanderWeele, 2019)

Example: Suppose p_{00}=0.1, p_{01}=0.2, p_{10}=0.4, then as p_{11} descends, we have:



Again, we can only unambiguously say the effect is larger if additive and multiplicative interaction are either <u>both</u> positive or <u>both</u> negative

Analogous classifications hold if both exposures are preventive, or if one is preventive and the other is causative (VanderWeele, 2019)

Conclusions

(1) The type of interaction analysis should be tied to the motivation for assessing interaction

(2) If we are simply assessing differential susceptibility both additive and multiplicative scales should be examined

(3) If we are interested in public health importance, additive scales should be used

(4) If we are interested in effects in different directions across populations we should assess qualitative interaction or use optimal subgroup selection

(5) If we are interested in mechanisms, assessing additive interaction can be informative for assessing sufficient cause interaction; multiplicative interaction may also give clues about mechanisms

(6) The absence of multiplicative interaction may aid generalizability

(7) Almost always, it is worth assessing interaction on both scales

Additional Slides

Case-Only Estimator for Interaction

Suppose also that G is independent of E in distribution in the population (this is plausible in many gene-environment interaction studies)

Suppose that data are only collected on the cases (Y=1), then the OR relating G and E among the cases is equal to the risk ratio multiplicative interaction: $\exp(\gamma_3)$ (Yang et al. 1999; cf. Piergorsch et al., 1994) – wow!

$$\frac{P(G=1|E=1,Y=1)/P(G=0|E=1,Y=1)}{P(G=1|E=0,Y=1)/P(G=0|E=0,Y=1)} = RR_{11}/(RR_{10} \times RR_{01})$$

This does, however, depend critically on the assumption that G and E are independent in distribution in the population and can result in severe bias if this is not the case (Albert et al., 2001)

Estimates and confidence intervals for the case-only estimator can be obtained by running a logistic regression of G on E and C among the cases:
$\text{logit}\{P(G=1|E=e,C=c,Y=1)\} = \theta_0 + \gamma_3 e + \theta_1' c$

Mechanistic Interaction: Epistasis

Somewhat related ideas and issues appear in the genetics literature and is sometimes referred to as “epistasis” (Bateson, 1909) or “compositional epistasis” (Philips, 2008; cf. Cordell, 2009; VanderWeele, 2010)

Often the term “epistasis” is used to describe statistical interaction between two genetic factors

Bateson (1909) used it for when one genetic variant would have an effect only when another variant was present

In our counterfactual notation, this would be: $Y_{11}=1$ but $Y_{10}=Y_{01}=Y_{00}=0$

Without monotonicity we can test this with RERI>2 (cf. VanderWeele, 2009 for further details)

Mechanistic Interaction: Biology and Models

- Thompson (1991) suggested that if an outcome required stages and one exposure acted on the first stage and another exposure acted on the second stage (a "multi-stage model") then if there were no biologic interaction, we would expect a multiplicative model.
- If the occurrence of a single adverse event was sufficient for the development of the disease (a "single-hit model") then in the absence of biologic interaction we would expect an additive model
- If the outcome occurred if an individual failed to experience any of one or more occurrences of a beneficial event (a "no-hit model", cf. Walter and Holford, 1978), then the model should again be multiplicative.

But all implications are one-way e.g. if an additive model fits well it is not necessarily the case that we have a "single-hit model" with no biologic interaction; it could equally be the case that we have a "multi-stage model" in which the factors operate antagonistically etc.

Motivation 5.
Attributing Effects to Interactions

If the exposures G and E are independent we can also decompose the total effect of one of the exposures (e.g. total effect of E) into two components:

$$(p_{e=1} - p_{e=0}) = (p_{01} - p_{00}) + (p_{11} - p_{10} - p_{01} + p_{00})P(G = 1)$$

(i) the component due to the second exposure E in the absence of the first, G
(ii) the component due to their interaction

Essentially the first component is the effect of E if we remove G
The second is how much of the effect of E we remove if we removed G

For binary exposures the proportion of the effect of E due to interaction is:

$$PAI_{G=0}(E) = \frac{(p_{11} - p_{10} - p_{01} + p_{00})P(G = 1)}{(p_{e=1} - p_{e=0})}.$$

VanderWeele and Tchetgen Tchetgen (2014) provide SAS and Stata code to do this automatically and handle more general cases and models

Attributing Effects to Interactions

If the exposures G and E are independent we could do this with both exposures; if the exposures are not independent we can stilll do this decomposition but only for the second exposure (e.g. if G affects E, we can still do this for E)

In the lung cancer example…
The proportion of the effect of E due to interaction is 36.6% $(95\%$ CI: $11.9\%, 61.3\%)$
The proportion of the effect of G due to interaction is 98.1% $(95\%$ CI: 66.1%, $129.9\%)$

Here G and E are not independent (G has a slight effect on E) so the interpretation of the second proportion becomes trickier but is probably a good approximation because the effect of G on E is very small

Policy Relevance: This approach of attributing a total effect to an interaction can help determine whether we might be able to eliminate the effect of an exposure by intervening on an effect modifier instead (note here we need the effects of both exposures to be unconfounded)