

Interaction and Effect Measure Modification

Mabel Carabali

EBOH, McGill University

Updated: 2024-10-23

Association between teacher distress and internalizing symptoms by elementary (n=14,665) and secondary students (n=8,903)

	Elementary	Secondary
	b (95% CI); p-value	b (95% CI); p-value
Teacher distress	0.08 (<-0.01, 0.16); p = 0.06	0.02 (-0.12, 0.16); p = 0.77
Student age (years)	0.27 (0.21, 0.34); p < 0.001	0.35 (0.27, 0.43); p < 0.001
Student gender: Female	1.43 (1.31, 1.56); p < 0.001	2.46 (2.26, 2.67); p < 0.001
Race or cultural group		
White (reference)		
East, Southeast, South Asian	-0.13 (-0.32, 0.07); p = 0.02	-0.33 (-0.65, -0.02); p = 0.04
Black	-0.76 (-1.06, -0.45); p < 0.00	-0.99 (-1.43, -0.56); p < 0.001
Other/Multiracial	0.007 (-0.17, 0.18); p = 0.60	0.29 (-0.02, 0.56); p = 0.04
Family assets	-0.08 (-0.39, 0.22); p = 0.54	-0.16 (-1.16, 0.68); p = 0.74
School SES2	-0.002 (-0.01, 0.01); p = 0.71	< 0.00 (<-0.01, 0.02); p = 0.45
School urbanity	-0.09 (-0.39, 0.22); p = 0.83	0.08 (-0.33, 0.48); p = 0.71
School safety	-0.44 (-0.47, -0.41); p < 0.001	-0.40 (-0.44, -0.35); p < 0.001
School safety*teacher distress	-0.01 (-0.01, <-0.00); p = 0.06	< 0.01 (<-0.01, 0.01); p = 0.97

Modified from: [The association between teacher distress and student mental health outcomes: a cross-sectional study using data from the school mental health survey](#). BMC Psychol 12, 583 (2024).

Expected competencies

- Basic knowledge about statistical interaction and Effect Measure Modification (EMM)
- Knows when to include an interaction (product) term
- Knows how to interpret an interaction term

Objectives

- Revise the concepts of EMM and Interaction.
- Revise the effect of interaction according to additive & multiplicative scales.
- Revise the main strategies for assessing and detecting EMM and interaction.

Effect Measure Modification (EMM) and Interaction

Most researchers seem to prefer a regression based approach to testing for *interaction* using the so called "interaction term" e.g. $(X \star M)$

- Yet assessing interaction is more than a product term in a regression model:
- There is statistical interaction (literally the result of the statistical test for the interaction term) and
- Causal interaction (whether intervening on the variables X and/or M will have a causal effect on outcome Y).

Often the terms interaction and effect modification are used interchangeably.

Biologic interaction.

In epidemiological discussions, sufficient-cause interaction is commonly referred to as biologic interaction (Rothman et al, 1980).

- Implying that, in biomedical applications, biological mechanisms could bring two exposures/treatments to act on each other to "cause" or produce an outcome.
 - Biological synergism is denoted when the two factors/exposures/treatments are the sufficient component cause of the outcome.
 - Both factors work together to produce the outcome. But it could be possible that these factors do not physically act on each other and thus that they do not interact in any biological sense.

For the rest of the lecture we will discuss EMM, Statistical and Causal Interaction

Effect Measure Modification (EMM) and Interaction

When we fit regression models with product terms, we have few possible interpretations:

- The coefficient for the product term could be interpreted as **EMM** .
- The coefficient for the product term could be interpreted as a measure of **interaction** .
- The coefficient for the product term could be interpreted as both **EMM** and **interaction** .
- The coefficient for the product term cannot be interpreted as **neither!**

On the Distinction Between Interaction and Effect Modification. VanderWeele, Tyler J. (2009)

Interaction and EMM

- **Interaction** refers to a situation whereby the effects of a exposure (covariate) of interest (on the outcome) and that of a third factor strengthen (synergism) or weaken each other (antagonism).
- **Effect modification** refers to the situation where the effect of a exposure (covariate) of interest differs depending on the presence or absence of a third factor (effect modifier). The third variable modifies the effect of the risk factor.
 - Synonyms: (Effect) heterogeneity¹

Both can be positive or negative; qualitative or quantitative.

¹ Side note on heterogeneity [Which of These Things Is Not Like the Others?](#)

Interaction and EMM

- **Effect *measure* modification**: examination of whether the effect of the risk factor is homogeneous or heterogeneous when stratified according to the suspected effect modifier.
 - Only interventions on the exposure variable of interest are considered
- **Interaction**: examination of whether the observed joint effect of the two variables of interest (exposure and third covariate) is the same or different than the expected from their independent effects.
 - Interventions on both, the exposure and third variable of interest are considered.

On the Distinction Between Interaction and Effect Modification. VanderWeele, Tyler J. (2009)

Forms of statistical interaction

According to the regression, the statistical interaction may be additive, multiplicative, or both.

- **Additive interaction** stipulates the summed effect of X and M will be greater than their individual effects,
 - In linear regression model, which is on the linear scale, the interaction term represents additive interaction.
 - The additive scale is a comparison of difference measures of association across strata (e.g., attributable risk/risk difference).
- **Multiplicative interaction** stipulates that the product of X and M will be greater than their individual effects.
 - In a GLM model (e.g., logistic or Poisson regression), which is on the log scale, the interaction term represents multiplicative interaction.
 - The multiplicative scale is a comparison of relative measures of association across strata (relative risk/odds ratio).

Additive or Multiplicative Interactions

As absolute measures are more frequently used in public health practice to demonstrate how interventions affect an outcome, it may be preferable to report additive interactions.

- But, the $X \star M$ term in linear regressions only indicate multiplicative interaction.
 - So, the additive component must be calculated separately, and depends in part on the parameterization of X and M

Fetal Sex and Race Modify the Predictors of Fetal Growth

Simone A. Reynolds · James M. Roberts ·
Lisa M. Bodnar · Catherine L. Haggerty ·
Ada O. Youk · Janet M. Catov

Published online: 17 July 2014
© Springer Science+Business Media New York 2014

Abstract The objective of this study is unknown if fetal sex and race modify the impact of maternal pre-pregnancy body mass index (BMI), and smoking on fetal growth. The authors studied markers of fetal growth in singleton offspring of 8,801 primiparous, normotensive women, enrolled in the Collaborative Perinatal Project. The authors tested for departures from additivity between sex/race and each predictor. The head-to-chest circumference ratio (HCC) decreased more, while birthweight and ponderal index (PI) increased more for each 1 kg/m² increase in pre-pregnancy BMI among term females versus males ($P = 0.07$, $P < 0.01$ and $P = 0.08$, interaction respectively). For term offspring of White compared with Black women, smoking independent of “dose” was associated with larger reductions in growth (165 g vs. 68 g reduction in birthweight, $P < 0.01$, interaction), greater reduction in

Statistical interaction terms between sex and each predictor and race and each predictor were used to test if the influence of maternal pre-pregnancy BMI and smoking varied by sex or race. Models with and without interaction terms were compared using the partial F test for linear models and the likelihood ratio test for logistic models ($P < 0.10$). It is proposed that departures from additivity (interaction in the additive scale) are more meaningful for interpreting the public health and biological significance than departures from multiplicativity (interaction on the multiplicative scale) [49–51]. Therefore departures from additive effects in the SGA models were also evaluated by calculating the synergy index (SI) (a test of interaction) [52] and its 95 % confidence interval [53]. The SI was calculated as the ratio between the combined effect and the sum of the individual effects of the risk factors [54].

Fetal Sex and Race Modify the Predictors of Fetal Growth

Table 1 Baseline characteristics of population of normotensive primiparous women

Characteristics	N = 8,801					
	Infant sex			Maternal race		
	Males (N = 4,597)	Females (N = 4,204)	P value	Blacks (N = 4,126)	Whites (N = 4,776)	P value
	n (%)	n (%)		n (%)	n (%)	
Maternal age						
<20	2,356 (52.0)	2,177 (48.0)	0.22	2,800 (61.8)	1,733 (38.2)	<0.01
20–29	2,125 (52.8)	1,898 (47.2)		1,218 (30.3)	2,805 (69.7)	
30+	116 (47.4)	129 (52.6)		79 (32.2)	166 (67.8)	
Maternal race						
White	2,483 (52.8)	2,221 (47.2)	0.27			
Black	2,114 (51.6)	1,983 (48.4)				
Smoking						
Nonsmoker	2,587 (52.5)	2,339 (47.5)	0.55	2,542 (51.6)	2,384 (48.4)	<0.01
Smoker	2,010 (51.9)	1,865 (48.1)		1,555 (40.1)	2,320 (59.9)	
Cigarettes/day						
Mean ± SD	6.04 ± 12.42	6.17 ± 12.8	0.63	3.99 ± 11.65	7.94 ± 13.09	<0.01
BMI (kg/m ²)						
Underweight (<18.5)	619 (53.8)	532 (46.2)	0.36	569 (49.4)	582 (50.6)	<0.01
Normal (18.5–24.9)	3,580 (52.3)	3,270 (47.7)		3,069 (44.8)	3,781 (55.2)	
Overweight (25.0–29.9)	330 (50.1)	329 (49.9)		374 (56.8)	285 (43.2)	
Obese (>30)	68 (48.2)	73 (51.8)		85 (60.3)	56 (39.7)	
Maternal SES score						
1 (low)	1,584 (51.4)	1,497 (48.6)	0.49	2,275(73.8)	806 (26.2)	<0.01
2 (mid)	1,238 (52.4)	1,126 (47.6)		1,236 (52.3)	1,128 (47.7)	
3 (high)	1,775 (52.9)	1,581 (47.1)		586 (17.5)	2,770 (82.5)	
Gestational age (completed weeks)						
Preterm <37	658 (14.3)	561 (13.3)	0.19	848 (20.7)	371 (7.9)	<0.01
Term ≥37	3,939 (85.7)	3,643 (86.7)		3,249 (79.3)	4,333 (92.1)	

Fetal Sex and Race Modify the Predictors of Fetal Growth

Table 4 Differential Influence of Smoking by Race

Fetal measures	Non Smoker Mean (SE) Term Black ^b	Smoker Mean (SE)	Black difference	Nonsmoker Mean (SE) Term White ^b	Smoker Mean (SE)	White difference	SMOKE × RACE Interaction P
Birthweight (g)	3,105.87 (9.52)	3,037.74 (11.99)	−68.13*	3,332.22 (9.13)	3,167.32 (8.96)	−164.90*	<0.01
Crown heel length (cm) ^a	49.92 (0.05)	49.44 (0.07)	−0.48*	50.66 (0.05)	49.98 (0.05)	−0.68*	0.06
Head circumference (cm) ^a	33.72 (0.03)	33.41 (0.04)	−0.31*	34.07 (0.03)	33.75 (0.03)	−0.32*	0.80
Chest circumference (cm) ^a	31.65 (0.04)	31.34 (0.05)	−0.31*	32.43 (0.04)	31.90 (0.04)	−0.52*	0.01
Placental weight (g) ^a	417.01 (2.28)	422.45 (2.75)	+5.44	434.47 (1.98)	438.09 (1.96)	+3.62	0.68
Ponderal index ^a	2.497 (0.007)	2.514 (0.008)	+0.017	2.562 (0.006)	2.536 (0.006)	−0.027*	<0.01
Head–chest circumference ratio ^a	1.067 (0.001)	1.068 (0.001)	+0.001	1.053 (0.001)	1.060 (0.001)	+0.007*	0.02
Fetal placental weight ratio ^a	7.627 (0.032)	7.348 (0.039)	−0.279*	7.836 (0.028)	7.377 (0.028)	−0.459*	<0.01
SGA n (%)	279 (13.6 %)	238 (19.5 %)	–	101 (4.5 %)	255 (12.0 %)	–	–
SGA (<10th percentile) ^c	Referent	OR 1.51 (1.24, 1.85)*	–	Referent	OR 2.90 (2.24, 3.77)*	–	<0.01 ^d
OR (95 % CI)							
SI (95 % CI) ^c		0.86 (0.67, 1.05)					0.89

Risk factors = Black race and smoking

* $P < 0.01$, † $P \leq 0.05$. Values are means and standard error unless otherwise specified. P values reported are for interaction in additive scale unless otherwise specified

Interaction terms : Example

```
library(MASS);
data(birthwt)
birthwt$smoke <- factor(birthwt$smoke, 0:1, c("non-smoker", "smoker"))
birthwt$race <- factor(birthwt$race, 1:3, c("white", "black", "other"))
birthwt$nonwhite <- birthwt$race != "white"
birthwt$nonwhite <- factor(as.numeric(birthwt$nonwhite), 0:1, c("white", "nonwhite"))
#head(birthwt[, c("bwt", "low", "smoke", "nonwhite", "age", "lwt")])
summary(birthwt[, c("bwt", "low", "smoke", "nonwhite", "age", "lwt")])
```

```
##           bwt           low           smoke           nonwhite           age
## Min.      : 709    Min.      :0.0000    non-smoker:115    white      :96    Min.      :14.00
## 1st Qu.:2414    1st Qu.:0.0000    smoker      : 74    nonwhite:93    1st Qu.:19.00
## Median :2977    Median :0.0000                                Median :23.00
## Mean     :2945    Mean     :0.3122                                Mean     :23.24
## 3rd Qu.:3487    3rd Qu.:1.0000                                3rd Qu.:26.00
## Max.     :4990    Max.      :1.0000                                Max.      :45.00
##           lwt
## Min.      : 80.0
## 1st Qu.:110.0
## Median :121.0
## Mean     :129.8
## 3rd Qu.:140.0
## Max.     :250.0
```

How do we interpret this regression output?

```
l13.mod1 <- lm(bwt ~ smoke * nonwhite, data = birthwt)
round(cbind(Beta =coef(l13.mod1), confint(l13.mod1)), 2)
```

##	Beta	2.5 %	97.5 %
## (Intercept)	3428.75	3226.09	3631.41
## smokesmoker	-601.90	-877.27	-326.54
## nonwhitenonwhite	-604.24	-862.17	-346.32
## smokesmoker:nonwhitenonwhite	419.49	-8.79	847.77

Which one is correct?

- On average a child born to white non-smoker parent weights 3428.75gs.
- On average a child born to non-white non-smoker parent weights -604.24g lower than from white parents who do not smoke.
- Non-white parents who do smoke have children with on average -786.65g lower birthweights than white parents who do not smoke.

Interpretation of the "interaction term"

(OLS: binary + binary covariates)

- **Non-smoking, white parents** Intercept, reference group, with an average birthweight given by the intercept: 3428.7g.
- **Smoking, white parent:** White parent who smoke have babies with on average -601.90g lower birthweights than white parent who do not smoke.
- **Non-smoking, non-white parent:** Non-white parent who do not smoke have babies with on average -604.2g lower birthweights than white parents who do not smoke.
- **Smoking, non-white parents:** Non-white parents who do smoke have babies with on average $(-601.90) + (-604.24) + 419.49 = -786.65\text{g}$ lower birthweights than white parents who do not smoke.

Interpretation of Interaction Coefficient: additional difference in means for non-reference levels of the two categorical variables.

How do we interpret this regression output?

```
l13.mod2 <- lm(bwt ~ smoke * age, data = birthwt)
round(cbind(Beta =coef(l13.mod2), confint(l13.mod2)), 2)
```

```
##              Beta    2.5 %  97.5 %
## (Intercept)  2406.06 1829.61 2982.51
## smokesmoker   798.17 -157.37 1753.72
## age           27.73   3.76   51.70
## smokesmoker:age -46.57 -86.91  -6.23
```

Which one is correct?

- On average a child born to a non-smoker parent weights 2406.06gs.
- The average birthweight decreases by 27.7grs per year of age of the parent.
- On average birthweight decreases 46.57 grs when the parent is a smoker.

Interpretation of the "interaction term"

(OLS: binary + continuous variable)

Without interaction: we assume that the slope of y over the continuous variable, x is the same regardless of the category of the variable, $m = 0$ or $m = 1$.

- Assume parallel regression lines for each group in m .

With an interaction term, we assume that the slope of y over x differs according to $m = 0$ or $m = 1$. Non parallel trend assumption.

- The **interaction coefficient** gives additional change in slope of y over x for the non-reference level of the nominal variable, $m = 1$.

The slopes are given by: $m = 0 : \beta_x; m = 1 : \beta_x + \beta_{x:m}$

Interpretation of the "interaction term"

(OLS: binary + continuous variable)

```
##              Beta    2.5 %   97.5 %  
## (Intercept)  2406.06 1829.61 2982.51  
## smoker      798.17 -157.37 1753.72  
## age         27.73   3.76   51.70  
## smoker:age  -46.57 -86.91  -6.23
```

- For non-smokers, average birthweight increases by 27.73g per year of age of the parent.
- For smokers, the average birthweight actually decreases by -18.84g ($27.73 + (-46.57)$) per one year increase in age of the parent.
- The mean difference between smokers and non-smokers *for age = 0* is 798.17grs ????

--

```
birthwt$agec <- birthwt$age - median(birthwt$age) #To center variable age, since age 0 does .  
l13.mod2c <- lm(bwt ~ smoke * agec, data = birthwt); #round(cbind(Beta =coef(l13.mod2c), con
```

How do we interpret this regression output?

```
birthwt$lwtc <- birthwt$lwt - median(birthwt$lwt) #To center variable lwtc, weight in pounds f  
l13.mod3 <- lm(bwt ~ agec * lwtc, data = birthwt)  
round(cbind(Beta =coef(l13.mod3), confint(l13.mod3)), 2)
```

```
##              Beta    2.5 %  97.5 %  
## (Intercept) 2912.11 2803.82 3020.40  
## agec        11.74   -9.59   33.06  
## lwtc         4.42    0.94    7.90  
## agec:lwtc   -0.30   -0.94    0.34
```

Which one is correct?

- Average birthweight increases by on average 11.74g for every year of the parent's age.
- Average birthweight increases by on average 4.42g for each pound of the parent's weight.
- Increasing age and weight of the parent make these associations slight less pronounced (-0.3g per year of age and pound).
- We expect $\text{agec} + \text{lwtc} = 16.16$
- We get $\text{agec} + \text{lwtc} + \text{agec:lwtc} = 15.86$

Interpretation of the "interaction term"

(OLS: continuous + continuous variable)

Including an **interaction term**, we assume that the slope of y over the continuous variable x_1 differs with respect to x_2 , and vice versa.

- Interpretation: the interaction term gives the change in slope of y over x_1 for each unit of x_2 , and the change in slope of y over x_2 for each unit of x_1 . The actual slopes are given by:
- Slope over x_1 : $\beta_{x_1} + x_2\beta_{x_1:x_2}$;
- Slope over x_2 : $\beta_{x_2} + x_1\beta_{x_1:x_2}$

Increasing age and weight of the parent make these associations slight less pronounced (-0.3g per year of age and pound).

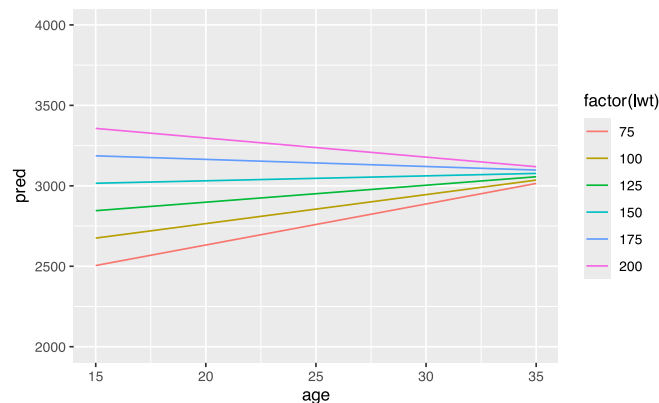
Interpretation of the "interaction term"

(OLS: continuous + continuous variable). Not always intuitive and this will change the calculation of the predicted values:

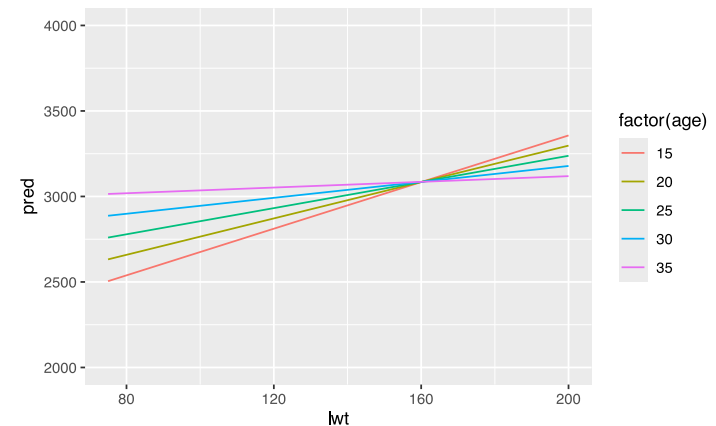
$$y = \beta_0 + \beta_1 agec_{(age-23)} + \beta_2 lwtc_{(lwt-121)} + \beta_3 agec * lwtc_{(age-23)(lwt-121)}$$

```
nd <- expand.grid(agec = seq(15, 35, 5) - 23, lwtc = seq(75, 200, 25) - 121)
nd$pred <- predict(l13.mod3, newdata = nd);
nd$age <- nd$agec + 23; nd$lwt <- nd$lwtc + 121
```

```
plot1<- qplot(age, pred, data = nd,
color = factor(lwt), geom = "line") +
ylim(2000, 4000)
plot1
```



```
plot2<- qplot(lwt, pred, data = nd,
color = factor(age), geom = "line") +
ylim(2000, 4000)
plot2
```



How do we interpret this regression output?

Logistic regression Outcome is lowbirthweight **low**, where yes=1 and no=0

```
l13.mod4 <- glm(low ~ smoke * nonwhite, data = birthwt, family = binomial)
```

##	Beta	2.5 %	97.5 %	##	OR	2.5 %	97.5 %
## (Intercept)	-2.3	-3.5	-1.4	## (Intercept)	0.1	0.0	0.2
## smokesmoker	1.8	0.7	3.1	## smokesmoker	5.8	1.9	21.4
## nonwhitenonwhite	1.7	0.6	3.0	## nonwhitenonwhite	5.4	1.9	19.6
## smokesmoker:nonwhitenonwhite	-1.1	-2.7	0.3	## smokesmoker:nonwhitenonwhite	0.3	0.1	1.4

- Smokers have 5.8 higher odds of having a low birthweight child compared to non-smokers.
- Non-white parents have a 5.4 higher odds of having low birthweight child compared to white parents.
- Non-white parents who smoke however have a 10 times higher odds of having a child with low birthweight than white parents who do not smoke.

Interpretation of the "interaction term"

(Logistic: binary + binary variable)

Including an interaction term, we assume that the OR comparing categories of x differs according to m , and viceversa.

- $OR < 1$ for the interaction indicates a less strong association than expected when considering them individually.
- $OR > 1$, association stronger than expected when considering them individually.

Interpretation: The interaction term gives multiplicative effect of non-reference levels of the two categorical variables. Examined by multiplying the odds ratios.

Interpretation of the "interaction term"

(Logistic: binary + binary variable)

To see the interaction effect of covariates x_1 and x_2 :

$OR_{x_1, x_2} = \exp(\beta_{x_1}) * \exp(\beta_{x_2}) * \exp(\beta_{x_1, x_2})$, equivalently, adding the coefficients and exponentiate them:

```
round(exp(coef(l13.mod4)["smokesmoker"]) * exp(coef(l13.mod4)["nonwhitenonwhite"]) * exp(coef
```

```
## smokesmoker  
##          10
```

```
round(exp(coef(l13.mod4)["smokesmoker"] + coef(l13.mod4)["nonwhitenonwhite"] + coef(l13.mod4
```

```
## smokesmoker  
##          10
```

```
5.76*5.43*0.32
```

```
## [1] 10.00858
```

How do we interpret this regression output?

##	OR	2.5 %	97.5 %
## (Intercept)	0.33	0.21	0.50
## smoker	2.05	1.09	3.89
## agec	0.92	0.84	1.00
## smoker:agec	1.08	0.95	1.23

- The odds of low birth weight child decreases by a factor of **0.92 per every year of the parent's among non-smokers**
- The odds of low birthweight child **decreases by a factor of 0.99 per every year of the parent's if they does smoke** (i.g., the additional change in odds of LBW for smokers) is **0.99**
- The odds of women who smoke AND increase a year of age **compared to non-smokers with average age is 2.03**
- The OR of low birthweight for **women who smoke, among women who increased one year is 2.2**

Interpretation of the "interaction term"

(Logistic: binary + continuous variable)

The interaction term gives additional change in odds for the non-reference level of the nominal variable, $m = 1$.

- The ORs are given by: $m = 0: e^{\beta_x}$ $m = 1: e^{\beta_x} e^{\beta_{x*m}}$
- The **additional** change $e^{\beta_x} e^{\beta_{x*m}}$ in odds (for one year increase) of low birthweight for smokers is 0.99

```
exp((-0.08288)+0.07308); #0.92*1.08 # this using the OR
```

```
## [1] 0.9902479
```

Interpretation of the "interaction term"

(Logistic: binary + continuous variable)

- The odds ratio for women who smoke **AND** increase a year of age $e^{\beta_x} e^{\beta_m} e^{\beta_{x*m}}$ compared to the non smokers and average age is 2.03

```
exp(0.71749+ (-0.08288)+0.07308); #alternatively using the ORs= 2.049*0.9204*1.075
```

```
## [1] 2.029298
```

Note: At least one exposure is <1.!

(Logistic: binary + continuous variable)

##	OR	2.5 %	97.5 %
## (Intercept)	0.33	0.21	0.50
## smokesmoker	2.05	1.09	3.89
## agec	0.92	0.84	1.00
## smokesmoker:agec	1.08	0.95	1.23

- The OR for women who smoke AND increase a year of age (**jointly exposed**) $e^{\beta_x} e^{\beta_m} e^{\beta_{x*m}}$ compared to the non smokers and average age is 2.03 ($2.05 * 0.92 * 1.08$)
- The OR on low birthweight for women who smoke, among women who increased one year $e^{\beta_x} e^{\beta_{x*m}}$ is 2.2 = ($2.05 * 1.08$)
- The **additional** change $e^{\beta_x} e^{\beta_{x*m}}$ in odds (for one year increase) of low birthweight for (among) smokers is 0.99 ($0.92 * 1.08$)

$$\text{logit}(Y) = \beta_0 + \beta_1 X + \beta_2 M + \beta_3 XM$$

In(odds)	M=0	M=1
X=0	β_0	$\beta_0 + \beta_2$
X=1	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$
In(OR)	M=0	M=1
X=0		
X=1	$(\beta_0 + \beta_1) - \beta_0 = \beta_1$	$(\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_2) = (\beta_1 + \beta_3)$
OR	M=0	M=1
X=0		
X=1	e^{β_1}	$e^{(\beta_1 + \beta_3)}$

- (`\beta_1`) will be either `smoking` or `"white vs non-white"` according to model specification.

(Logistic: binary + continuous variable) $logit(Y) = \beta_0 + \beta_1 X + \beta_2 M + \beta_3 XM$

Here the OR_{00} is the reference, derived from the intercept (interpreted as the average log(odds) or odds) We can obtain a table with all combinations using the `interactionR` Package

```
table_objectlbw = interactionR(l13.mod5, exposure_names = c("smokesmoker", "agec"),
                              ci.type = "mover", ci.level = 0.95, em = F, recode = F)
kable(table_objectlbw$dframe[1:9,], digits = 2)
```

Measures	Estimates	CI.l	CI.u	p
OR00	1.00			
OR01	0.92	0.84	1.01	0.07
OR10	2.05	1.09	3.87	0.03
OR11	2.03	1.07	3.86	0.03
OR(agec on outcome [smokesmoker==0])	0.92	0.84	1.01	0.07
OR(agec on outcome [smokesmoker==1])	0.99	0.90	1.09	0.85
OR(smokesmoker on outcome [agec==0])	2.05	1.09	3.87	0.03
OR(smokesmoker on outcome [agec==1])	2.20	1.14	4.25	0.02
Multiplicative scale	1.08	0.95	1.22	0.26

Interaction & EMM in observational settings

Let's think about a randomized experiment

In a randomized trial setting, “interaction” is a joint causal effect, which we obtain by factorial design.

- We randomized X & M , and look at their causal effects separately and together.

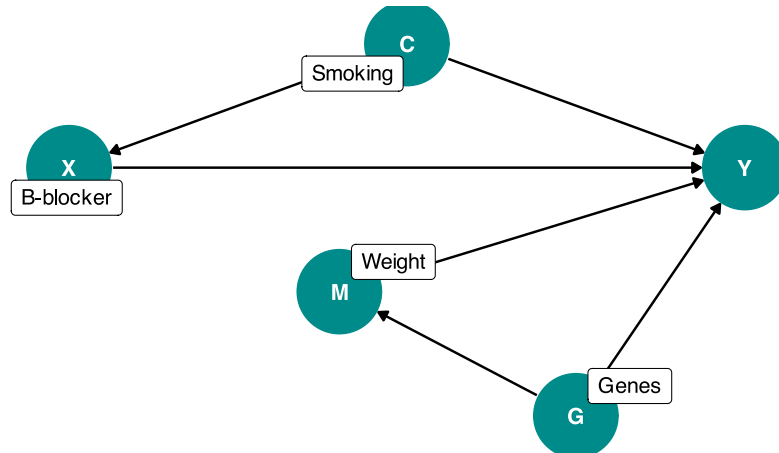
In a randomized setting, “effect measure modification” is a subgroup analysis.

- We randomize X , and look at the effect of X where $M=1$ and separately where $M=0$.

Epidemiology by design by Daniel Westreich

Consider this DAG

DAG for EMM and Interaction

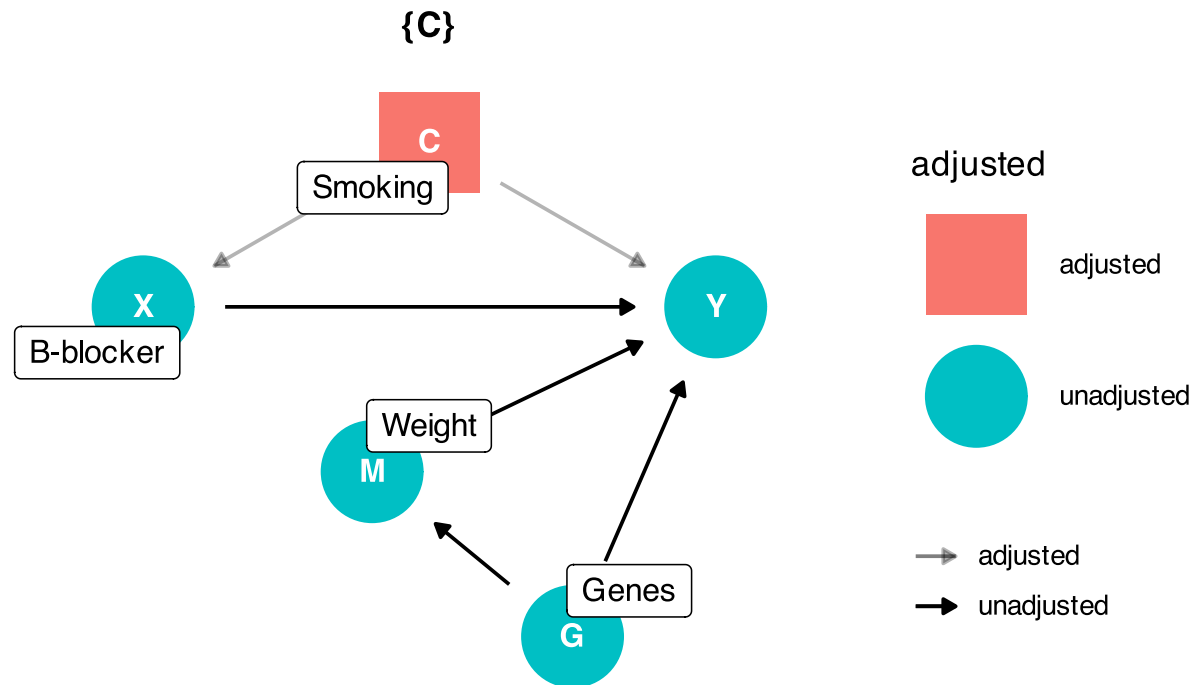


What is the effect of Beta-blocker on Y (Heart attack)?

- Assume treatment variation irrelevance.
- Assume this DAG is correct (! Big assumption). Open backdoor paths are:
 $\beta - blocker \leftarrow Smoking \rightarrow Heartattack$

- We entirely ignore weight/genes/heart attack pathway right now – it's not relevant to the beta-blocker-heart attack relationship.
- Controlling for smoking will block open backdoor path.

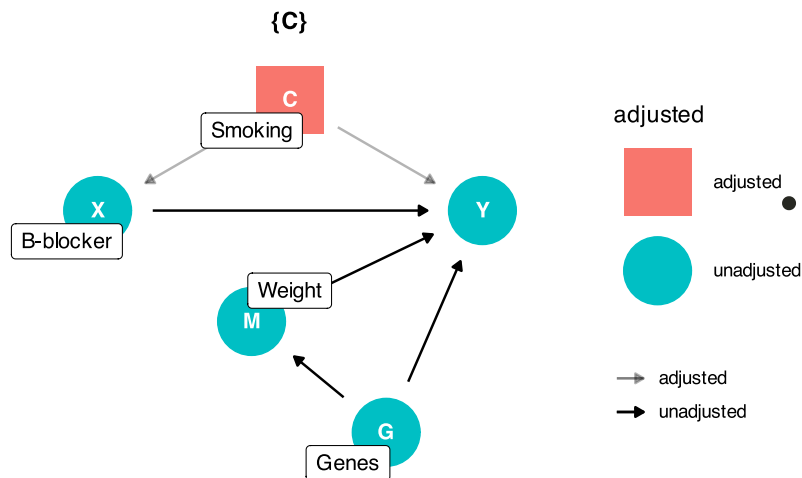
Interaction and EMM



Controlling for smoking will block open backdoor path. Assume that we can now assess a causal effect of Beta-blocker on Heart attack.

- Can we, from this DAG, obtain an **unconfounded** effect of weight on heart attack?

Interaction and EMM



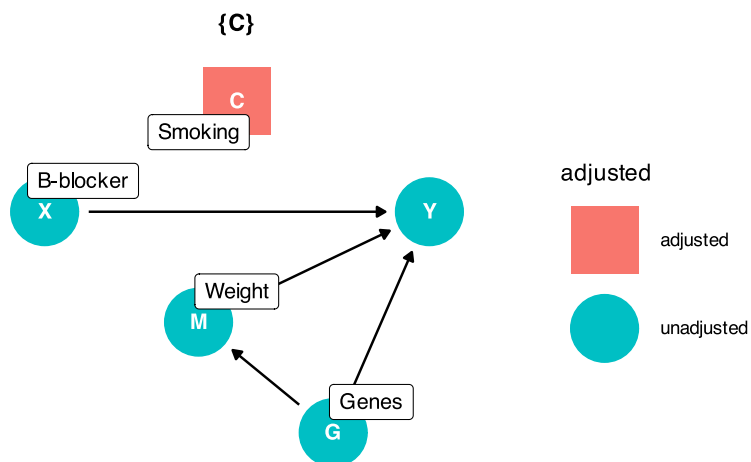
From the present DAG, we can examine effect measure modification of Beta-blockers on heart attack by Weight.

- Remember, effect measure modification is asking about impact of Beta-blockers by **observed** weight.

If weight is high/low, could just ask: among those with high weight, what's the causal effect of beta-blockers on heart attack. And, among those with low weight...

Can we, from this DAG, obtain an *unconfounded* effect of weight on heart attack?

Can we, from this DAG, obtain an unconfounded effect of weight on heart attack?



- No: confounding by genes.

If we wanted an **INTERACTION** of Beta-blockers and weight, then, we'd need to control for confounding of the weight-heart attack relationship by genes.

Different model.

VanderWeele T. On the distinction between Interaction and Effect Modification. Epidemiology 2009.

Interaction and EMM in cohorts

Translating concepts from a trial... outcome **Y**.

- Think of your main exposure **X** as the “randomized” exposure.
 - You try to control for confounding to make it as close to “randomized” as you can.
There are statistical subtleties here to worry about later (“collapsibility”)

Then think of the second variable **M**: are you controlling for confounders of the $M \rightarrow Y$ relationship?

- **Yes:** then you’re estimating **interaction** (because now you’re trying to make the second variable as close as possible to “randomized” – as in a **factorial design**)
- **No:** then you’re estimating **effect measure modification** (subgroup analysis)

Epidemiology by design by Daniel Westreich

Interaction vs. effect measure modification

Broadly, both can be assessed in observational settings.

In the same way that interaction in a randomized trial is a factorial trial,

- Interaction in an observational study means estimating two causal effects,
 - and meeting causal identification conditions for **BOTH**.

Interaction vs. EMM under counterfactuals

- A variable M is a modifier of the effect of X on Y when the average causal effect of X on Y varies across levels of M .
 - The concept of EMM refers to the causal effect of X , **not to the causal effect of M** .
- When we say that M modifies the effect of X we are **not considering** M and X as variables of equal status, because **only X is considered** to be a variable on which we could hypothetically intervene.
- EMM involves the counterfactual outcomes Y^a , **not the counterfactual** outcomes $Y^{x,m}$.
- Interaction between X and M gives **equal status** to both treatments X and M .
 - It refers to the joint causal effect of two treatments X and M , and thus involves the counterfactual outcomes $Y^{x,m}$ under a joint intervention.

Notes on Effect *Measure* Modification

From: Hernan & Robins book, What if?

and some **worked** examples.

H&R What if? 4.2 Stratification to identify EMM

To identify effect modification by M in an ideal experiment with unconditional randomization, one just needs to conduct a stratified analysis by computing the association measure in each level of M . (Average causal effects in subsets of the population).

Read the **H&R What if? Fine Point 4.1: Effect in the treated ($X=1$)** for additional details on EMM estimation.

H&R What if? 4.3 Why care about EMM

The extrapolation of causal effects computed in one population to a second population, i.e., transportability of causal inferences across populations.

What factors affect generalizability? (See Fine Point 4.2.)

- Evaluating the presence of EMM is helpful to identify the subgroups in the population that would benefit most from an intervention.
- Additive, but not multiplicative, EMM is the appropriate scale to identify the groups that will benefit most from intervention. **Why do we report ratios?** H&R What If? propose: The identification of effect modification may help understand the biological, social, or other mechanisms leading to the outcome. Koopman JS, Weed DL. Epigenesis theory: a mathematical model relating causal concepts of pathogenesis in individuals to disease patterns in populations. Am J Epidemiol. 1990

But also refuted, for example by: Thompson WD. Effect modification and the limits of biologic inference from epidemiologic data. Journal of Clinical Epidemiology 1991;44:33-6/pdf)

H&R What if?

4.4 Stratification as a form of adjustment

- Adjustment for **M** is typically accomplished by stratifying estimation on categories of **M** (no confounding by **M** within categories).
- Standardization is different, because the categories are weighted, but there is no conditioning.
 - It is a control method that does not require homogeneity.
 - Conditional versus marginal estimates are different for non-collapsible measures (OR, IRR, HR, etc). See [Fine Point 4.3](#).

Qualitative vs. Quantitative EMM and Interaction

- When the association between factor **X** and outcome **Y** exists and is of the same direction in each stratum formed by **M**, but the strength of the association varies across strata, we call this quantitative effect measure modification.
 - Example: $M=1 \rightarrow RR = 5.7, 95\%CI(4.3, 7.4)$ & $M=0 \rightarrow RR = 1.8, 95\%CI(1.1, 2.4)$
- However, when the effects of **X** on **Y** changes in direction (crossover) within levels of third variable **M**, we call the qualitative effect measure modification.
 - Example: $M=1 \rightarrow RR = 5.7, 95\%CI(4.3, 7.4)$ & $M=0 \rightarrow RR = 0.6, 95\%CI(0.3, 0.8)$
- Qualitative EMM also may be said to occur when the association between **X** and **Y** disappears in some levels of **M**. It refers to a genuine change in interpretation of the exposure disease relationship over levels of **M**.

Low birthweigh example

X = Smoking (can intervene); **M** = white vs non-white (cannot/ *don't want* to intervene).

Interaction term regression

```
##                                OR 2.5 % 97.5 %
## (Intercept)                  0.10  0.03  0.25
## smoker                       5.76  1.94 21.37
## nonwhitenonwhite             5.43  1.91 19.63
## smoker:nonwhitenonwhite      0.32  0.06  1.39
```

White strata

```
l13.lbwwhite<- glm(low ~ smoke, family = bi
                  data = subset(birthwt, n
round(cbind("OR" = exp(coef(l13.lbwwhite))),
```

```
##                                OR 2.5 % 97.5 %
## (Intercept) 0.10  0.03  0.25
## smoker      5.76  1.94 21.37
```

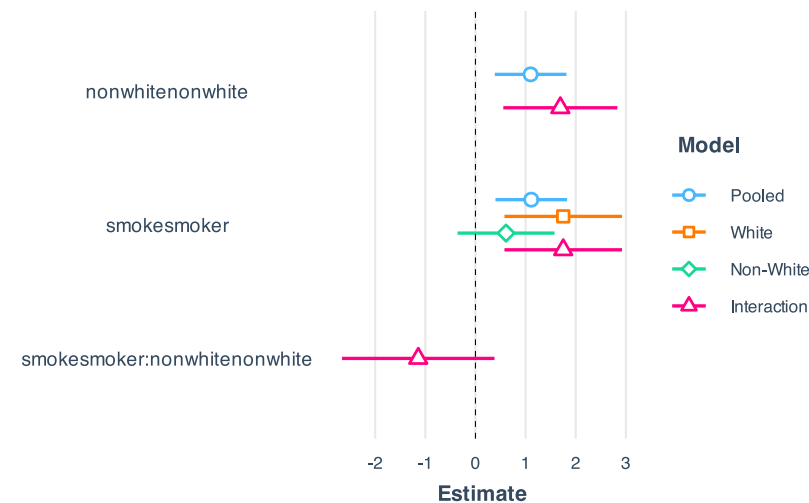
Non-white strata

```
l13.lbwnonwhite<- glm(low ~ smoke, family =
                    data = subset(birthwt
round(cbind("OR" = exp(coef(l13.lbwnonwhite
```

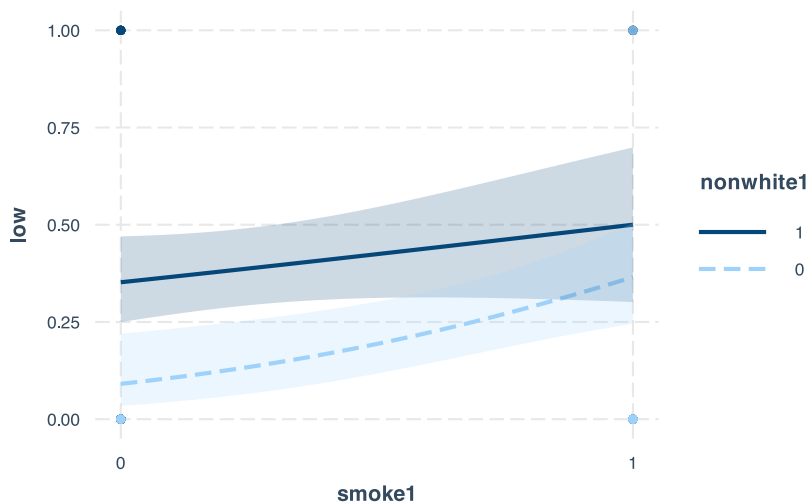
```
##                                OR 2.5 % 97.5 %
## (Intercept) 0.54  0.33  0.88
## smoker      1.84  0.70  4.90
```

Low birthweigh example

Coefficients from the regression



EMM Plot - M is "White" vs "Non-white"



Low birthweigh example: **X** = white vs non-white; **M** = Smoking: **Illustation only** or descriptive analysis -No causal interpretation!

Pooled, "adjusted" regression

```
l13.lbwsmkpooled<- glm(low ~ nonwhite+ smok  
round(cbind("OR" = exp(coef(l13.lbwsmkpooled
```

##		OR	2.5 %	97.5 %
##	(Intercept)	0.16	0.08	0.30
##	nonwhitenonwhite	3.01	1.50	6.30
##	smokesmoker	3.04	1.51	6.35

White strata

```
l13.lbwsmk0<- glm(low ~ nonwhite, family =  
round(cbind("OR" = exp(coef(l13.lbwsmk0))),
```

##		OR	2.5 %	97.5 %
##	(Intercept)	0.10	0.03	0.25
##	nonwhitenonwhite	5.43	1.91	19.63

Interaction term

```
l13.mod4 <- glm(low ~ smoke * nonwhite, dat  
round(cbind("OR" = exp(coef(l13.mod4))), exp
```

##		OR	2.5 %	97.5 %
##	(Intercept)	0.10	0.03	0.25
##	smokesmoker	5.76	1.94	21.37
##	nonwhitenonwhite	5.43	1.91	19.63
##	smokesmoker:nonwhitenonwhite	0.32	0.06	1.39

Non-white strata

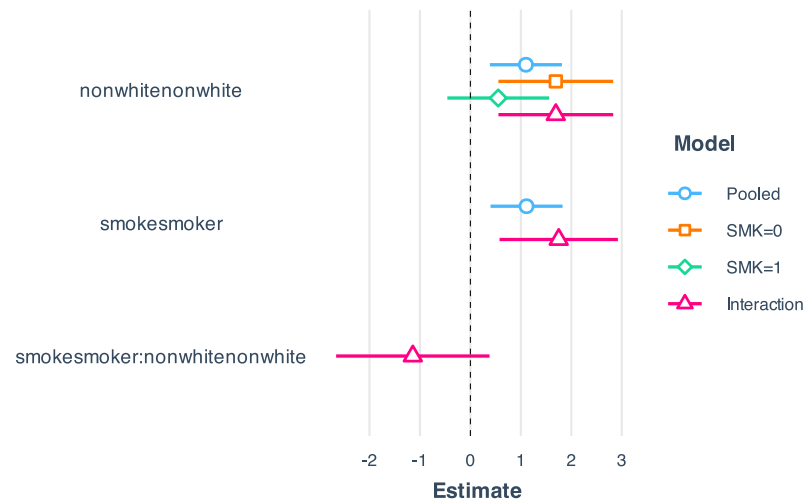
```
l13.lbwsmk1<- glm(low ~ nonwhite, family =  
round(cbind("OR" = exp(coef(l13.lbwsmk1))),
```

##		OR	2.5 %	97.5 %
##	(Intercept)	0.58	0.32	1.00
##	nonwhitenonwhite	1.74	0.63	4.82

Low birthweigh example; **X** = white vs non-white; **M** = Smoking

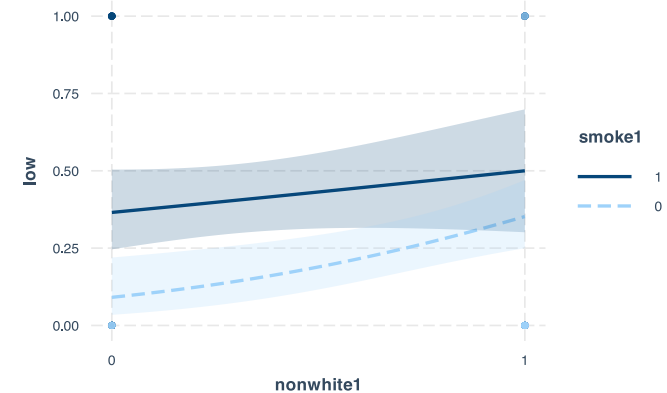
Coefficients from the regression

```
plot_summs(l13.lbwsmkpooled, l13.lbwsmk0, l13.lbwsmk1,
  robust = list(FALSE, FALSE, FALSE),
  model.names = c("Pooled", "SMK=0", "SMK=1", "Interaction"))
```



EMM Plot - **M** is Smoking

```
interact_plot(l13.mod4a, pred = nonwhite1,
  modx = smoke1,
  interval = TRUE,
  plot.points = TRUE)
```



How many people do we have here?

189, 15

Low birthweigh example; smoking & white vs non-white comparisons

Estimated ORs and contrasts.

Measures	Estimates	CI.l	CI.ul	p
OR00	1.00			
OR01	5.43	1.74	16.95	0.00
OR10	5.76	1.78	18.60	0.00
OR11	10.00	2.66	37.60	0.00
OR(nonwhitenonwhite on outcome [smokesmoker==0])	5.43	1.74	16.95	0.00
OR(nonwhitenonwhite on outcome [smokesmoker==1])	1.74	0.63	4.76	0.29
OR(smokesmoker on outcome [nonwhitenonwhite==0])	5.76	1.78	18.60	0.00
OR(smokesmoker on outcome [nonwhitenonwhite==1])	1.84	0.70	4.84	0.22
Multiplicative scale	0.32	0.07	1.46	0.14

When an interaction term is **Statistically significant**?

Interaction tests generally have about 25% of the power of main effects tests (Greenland 1983).

- Need **4 times the sample size** to estimate an interaction that is the same size as the main effect
 - To be able to estimate the interaction to the same level of accuracy as the main effect, we would need four times the sample size.
- **Need 16 times** the sample size to estimate an **interaction that is half the size as the main effect**.

Using a higher Type I error criterion, e.g., 0.15 or 0.2 makes it easier to detect real heterogeneity, but raises the % of times that heterogeneity will be declared erroneously due to Type I error.

*"...interactions can be important; we just need **to accept that in many settings we won't be able to attain anything like near-certainty regarding the magnitude or even direction of particular interactions**. It is typically not appropriate to aim for "statistical significance" or 95% intervals that exclude zero, and it often will be appropriate to use prior information to get more stable and reasonable estimates, and to accept uncertainty, not acting as if interactions of interest are zero just because their estimate is not statistically significant".* ROS-Gelman, Hill & Vehtari

H&R What if? 4.6 EMM and adjustment methods

H&R, What if? Heart transplant example computed four causal risk ratios and obtained four different numbers: 0.8, 2.0, 0.5 and 1.0.

- [Standardization and IP weighting](#) yield the average causal effect in the entire population = 0.8.
- [Stratification](#) yields the conditional causal risk ratios 2.0 and 0.5.
- [Matching](#), using matched pairs described in the text gives 1.0.

All of these numbers are “correct”.

- This example highlights the primary importance of specifying the population, or the subset of a population, to which the effect measure corresponds.
- [Technical Point 4.2](#): Conditional estimation assumes homogeneity across strata (e.g. regression models, Mantel-Haenszel estimator, etc).

EMM - Analysis

When dealing with stratum-specific effect measures, one commonly used strategy to reduce the variability of the estimates is to combine all stratum-specific effect measures into one pooled stratum-specific effect measure.

- If the effect measure is the same in all strata (no EMM), then the pooled effect measure will be a more precise estimate of the common effect measure.
- Several methods (e.g., Mantel-Haenszel) yield a pooled estimate, sometimes by computing a weighted average of the stratum-specific effect measures with weights chosen to reduce the variability of the pooled estimate.
- Pooled effect measures can also be computed using regression models that include **all possible product terms** between all covariates **M**, but no product terms between treatment **X** and covariates **M**, i.e., models “saturated” with respect to **M**.

H&R What if? EMM - Analytic considerations

- The main goal of pooling is precision of the common stratum-specific effect measure, but the pooled effect measure is still a **conditional effect measure**!

In Hernan's heart transplant example, the pooled stratum-specific RR by the Mantel- Haenszel method) was 0.88.

- This result is only meaningful if the stratum-specific RR's 2 and 0.5 are indeed estimates of the same stratum-specific causal effect.
 - Additive effect modification is the appropriate scale to identify the groups that will benefit most from intervention.
- In the absence of additive effect modification, learning that there is multiplicative effect modification may not be very helpful for decision making.

Interaction and EMM

- Interaction is defined in terms of the effects of **two (2) interventions** while EMM is defined in terms of the **effect of one intervention varying across strata of a second variable**.
- EMM can be present with no interaction; interaction can be present with no EMM.
- There are settings in which it is possible to assess effect modification but not interaction, or to assess interaction but not effect modification.

EFFECT MODIFICATION AND INTERACTION

"In many settings, of course, interaction and effect modification will coincide. The question thus arises whether we can characterize settings in which interaction and effect modification do indeed coincide. A characterization can be given in terms of causal DAGs. It follows from the rules of causal DAGs (see Rule 2 of Pearl's do-calculus⁸) that we will have that interaction and effect modification will coincide (because $E[Deq] = E[De \mid Q = q]$), whenever all paths between D and Q are blocked by E on a graph obtained by modifying the original graph to remove the arrows into E and the arrows emerging from Q ."

On the Distinction Between Interaction and Effect Modification. VanderWeele, Tyler J. (2009)

EMM or Interaction?



Virgin "Cuba-libre" - The Big Bang Theory

QUESTIONS?

COMMENTS?

RECOMMENDATIONS?

EMM , Interactions and DAGs

H&R What if? 4.2 Stratification to identify EMM

To identify effect modification by **M** in an ideal experiment with unconditional randomization, one just needs to conduct a stratified analysis by computing the association measure in each level of **M**. (Average causal effects in subsets of the population).

H&R What if? Fine Point 4.1: Effect in the treated ($X=1$).

- 1) The ATT is not null if $Pr[Y^{x=1} = 1|X = 1] \neq Pr[Y^{x=0} = 1|X = 1]$ or, by consistency, if $Pr[Y = 1|X = 1] \neq Pr[Y^{x=0} = 1|X = 1]$.
- 2) There is a causal effect in the treated if the observed risk among the treated individuals does not equal the counterfactual risk had the treated individuals been untreated.
- 3) **ATT** will differ from the average effect in the population if the distribution of individual causal effects varies between the treated and the untreated.
- 4) Epidemiologists refer to the ATT ratio (causal risk ratio in the treated) as the “Standardized Mortality/Morbidity Ratio” (SMR): $Pr[Y = 1|X = 1]/Pr[Y^{x=0} = 1|X = 1]$.
- 5) Hypothetically, there is an analogous causal effect in the untreated (ATU), but this is uncommon in both epidemiology and economics (although it is occasionally used in both fields).

H&R What if?

4.4 Stratification as a form of adjustment

- Adjustment for **M** is typically accomplished by stratifying estimation on categories of **M** (no confounding by **M** within categories).
- Standardization is different, because the categories are weighted, but there is no conditioning.
 - It is a control method that does not require homogeneity.
 - Conditional versus marginal estimates are different for non-collapsible measures (OR, IRR, HR, etc). See [Fine Point 4.3](#).

4.5 Matching as another form of adjustment

- NB: Matching used for adjustment in cohort studies, but NOT in case-control studies. In case-control studies it is used to gain efficiency. Matching actually CAUSES confounding in case-control studies, and therefore, one must adjust for matching factors.
 - Matching ensures positivity in the matched population because strata with only treated, or untreated, individuals are excluded from the analysis. Any target population can be chosen.**ATT**

EFFECT MODIFICATION WITH NO INTERACTION AND INTERACTION WITH NO EFFECT MODIFICATION

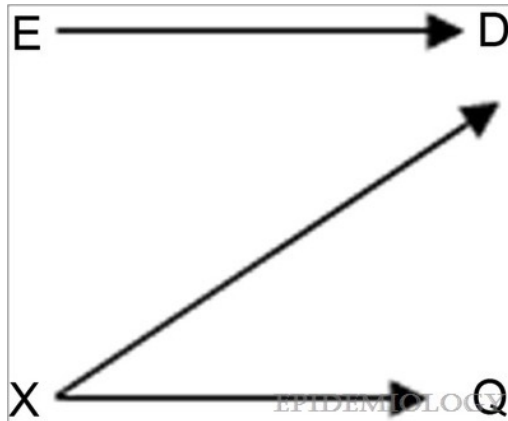


FIGURE 1.: Effect modification by Q of the effect of E on D without interaction between the effects of E and Q on D.

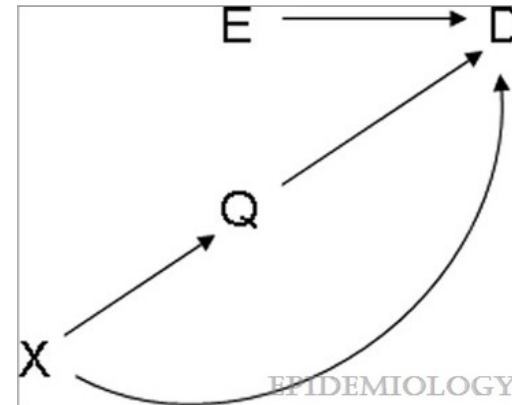


FIGURE 2.: Potential interaction between the effects of E and Q on D without effect modification by Q of the effect of E on D.

On the Distinction Between Interaction and Effect Modification. VanderWeele, Tyler J. (2009)

EFFECT MODIFICATION WITH NO INTERACTION AND INTERACTION WITH NO EFFECT MODIFICATION

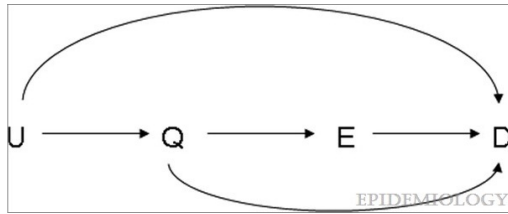


FIGURE 3.: Identification of effect modification of the effect of E on D by Q without identification of the joint effects of E and D.

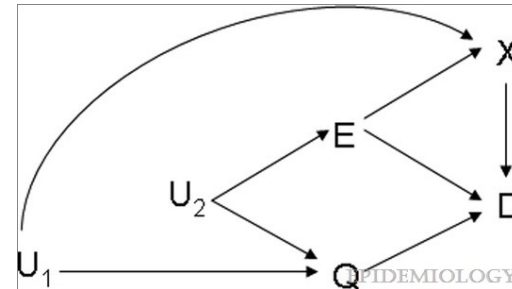


FIGURE 4.: Identification of the joint effects of E and D without identification of effect modification of the effect of E on D by Q.

On the Distinction Between Interaction and Effect Modification. VanderWeele, Tyler J. (2009)