Are you looking for the right interactions?

Additive versus multiplicative interactions with dichotomous outcome variables

Melanie M. Wall
Departments of Psychiatry and Biostatistics
New York State Psychiatric Institute and Mailman School of Public Health
Columbia University
mmwall@columbia.edu

This presentation Slides 1-17 given at the ENAR Biometrics meeting in Orlando Florida March 2013
Starting on slide 18 is a similar presentation given to the Epi department at Columbia University in 2012 with a different demonstration of software for estimation

Background

- The meaning of the term “interaction” can be cause for confusion
- In statistical terms, an interaction is present when the effect of one variable on the outcome depends on the levels of another variable.
- **Problem:** Whether a statistical interaction is found or not depends on how effects are measured, i.e. depends upon the scale (additive or multiplicative).
- **This is well known in the epidemiology literature, but not well (enough) know among biostatisticians**

Classic psychiatry dataset sparking debate about additive vs multiplicative interaction (1978)

		Outcome (Depression)			
Prior Vulnerability	Exposure Stress	No	Yes	Risk of Depression	
No	No	191	2	0.010	P00
	Yes	79	9	0.102	P01
Effect of Stress given No Vulnerability ->		OR = 10.9 (2.3, 51.5) RR = 9.9 (2.2, 44.7)		RD = 0.092 (0.027,0.157)	
Yes	No	60	2	0.032	P10
	Yes	52	24	0.316	P11
Effect of Stress given Vulnerability ->		OR = 13.8 (3.1,61.4) RR = 9.8 (2.4, 39.8)		RD=0.284 (0.170,0.397)	

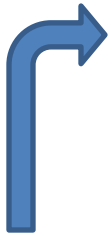
OR = Odds Ratio (95% Confidence Interval) <-compare to 1

RR = Risk Ratio (95% Confidence Interval) <-compare to 1

RD = Risk Difference (95% Confidence Interval) <-compare to 0

Does Vulnerability Modify the Effect of Stress on Depression?

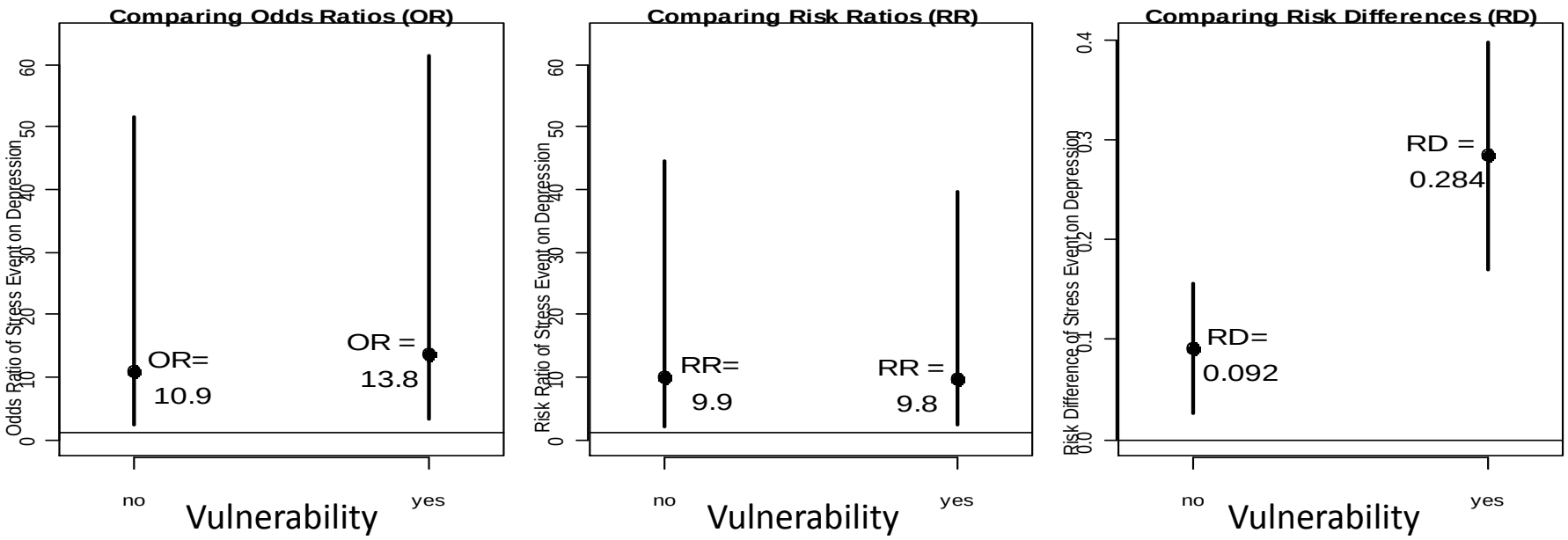
- On the **multiplicative Odds Ratio scale**, is 10.9 sig different from 13.8?
  - Test whether the ratio of the odds ratios (i.e. 13.8/10.9 = 1.27) is significantly different from 1.
- On the **multiplicative Risk Ratio scale**, is 9.9 sig different from 9.8?
  - Test whether the ratio of the risk ratios (i.e. 9.8/9.9 = 0.99) is significantly different from 1.
- On the **additive Risk Difference scale**, is 0.092 sig different from 0.284?
  - Test whether the difference in the risk differences (i.e. 0.28-0.09 = 0.19) is significantly different from 0.



Rothman calls this difference in the risk differences the "interaction contrast (IC)"

IC = (P11 - P10) – (P01 - P00)

Comparing stress effects across vulnerability groups
Different conclusions on multiplicative vs additive scale



95% confidence intervals for Odds Ratios overlap

-> no statistically significant multiplicative interaction OR scale

95% confidence intervals for Risk Ratios overlap

-> no statistically significant multiplicative interaction RR scale

95% confidence intervals for Risk Differences do not overlap

-> statistically significant additive interaction

Test for multiplicative interaction on the OR scale- Logistic Regression with a cross-product

IN SAS:

```
proc logistic data = brownharris descending;
model depressn = stressevent   vulnerability   stressevent*vulnerability;
run;
```

Analysis of Maximum Likelihood Estimates

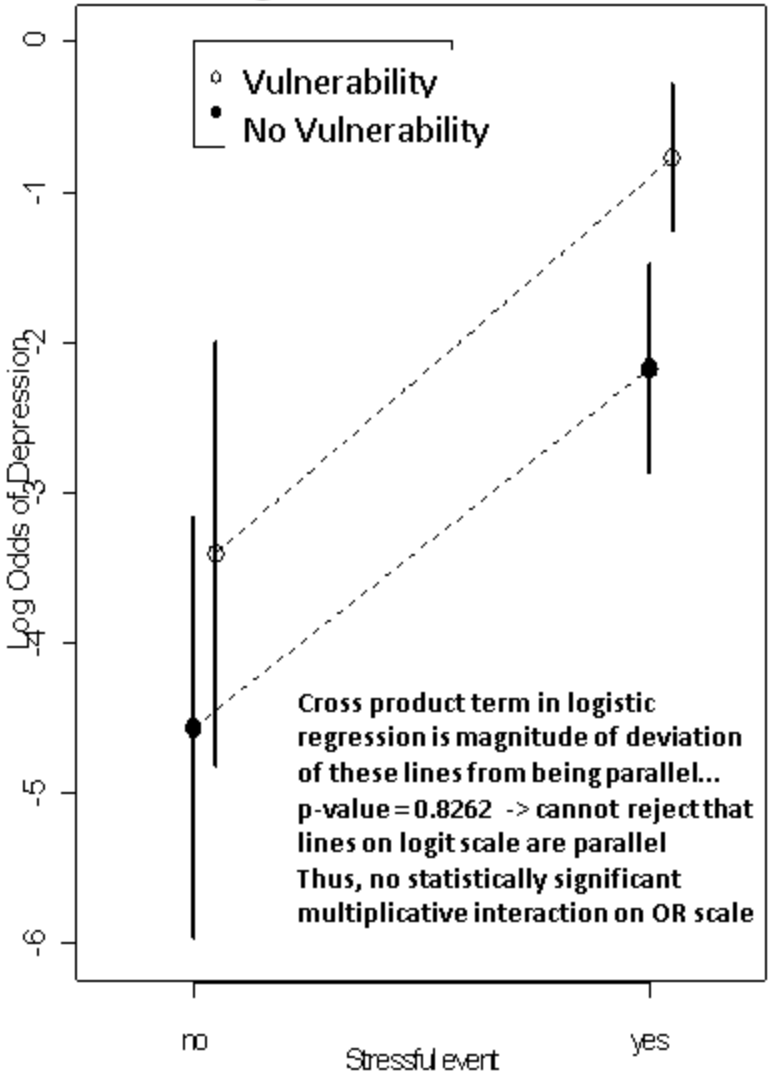
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.5591	0.7108	41.1409	<.0001
stressevent	1	2.3869	0.7931	9.0576	0.0026
vulnerability	1	1.1579	1.0109	1.3120	0.2520
stresseve* vulnerabi	**1**	**0.2411**	**1.0984**	**0.0482**	**0.8262**

exp(.2411) = 1.27 =
Ratio of Odds ratios =13.846/10.880
Not significantly different from 1

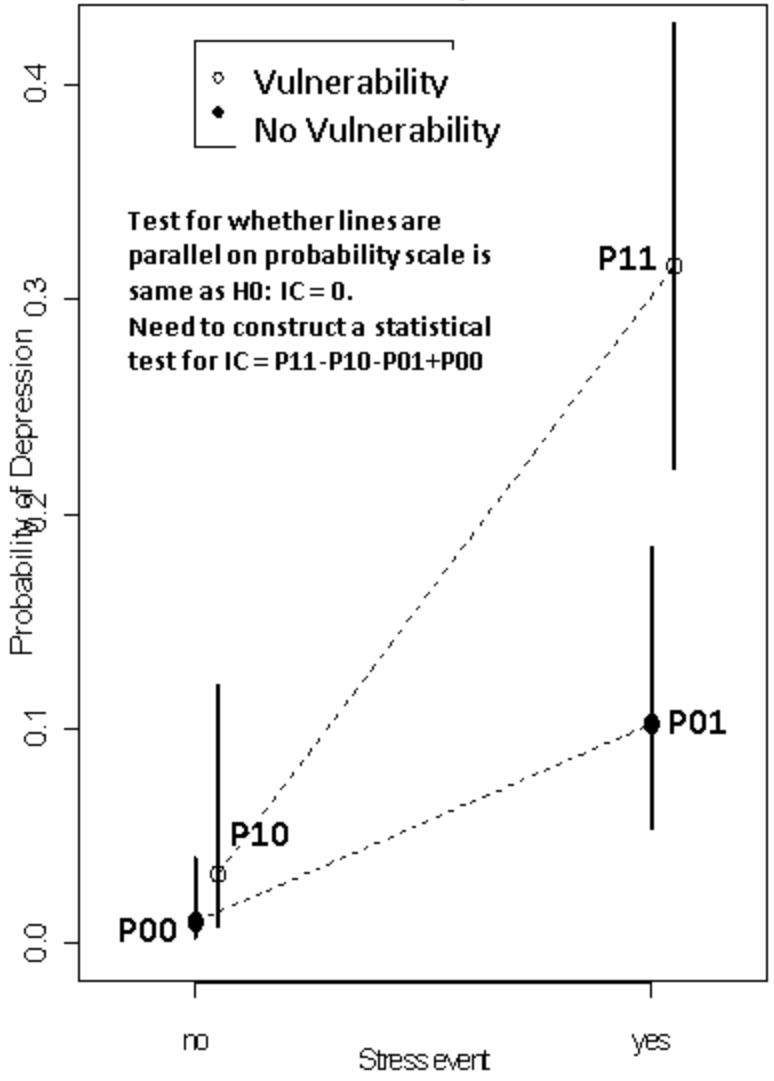
"multiplicative interaction" on OR scale is not significant

Test for interaction: Are the lines Parallel?

Log Odds scale



Probability scale



Test for additive interaction on the probability scale- 3 Different Strategies

The Additive Interaction Contrast (IC) is the Difference of Risk Differences:

IC = (P11 - P10) – (P01 - P00) = P11-P10-P01+P00. We want to test this = 0.

1. Directly fit a **linear risk model**:

Risk = b0 + b1 * STRESS + b2 * VULN + b3*STRESS*VULN; **b3 = IC**.

2. Fit a **logistic regression model then back-transform** to get probabilities (P11, P10, P01, P00) to estimate and test IC.

3. Divide the IC by P00 and get a contrast of risk ratios:

IC Ratio = RR(11) – RR(10) – RR(01) + 1

Estimate Relative Excess Risk due to Interaction (RERI),

test RERI= 0

Strategy #1 Using a linear risk model

(A)linear binomial model

(B)linear normal model using robust standard errors

PROS:

- Contrast of interest is directly estimated and tested
- Covariates easily included
- Can be used with continuous predictors in the interaction
- Can do in most any statistical software

CONS:

- Linear model for probabilities can be greater than 1 and less than 0 (outside the parameter space)
- Convergence problems common for the linear binomial model using maximum likelihood estimation

Testing for additive interaction on the probability scale
Strategy #1a: Use linear binomial regression with a cross-product

link=identity dist=binomial tells SAS to do linear binomial regression. lrci outputs likelihood ratio (profile likelihood) confidence intervals.

IN SAS:

```
proc genmod data = individual descending;
model depressn = stressevent   vulnerability   stressevent*vulnerability/
link = identity dist = binomial lrci;
run;
```

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Likelihood Ratio 95% Confidence Limits		Wald Chi-Square	Pr>ChiSq
Intercept	1	0.0104	0.0073	0.0017	0.0317	2.02	0.1551
stressevent	1	0.0919	0.0331	0.0368	0.1675	7.70	0.0055
vulnerability	1	0.0219	0.0236	-0.0139	0.0870	0.86	0.3534
stresseve*vulnerabil	**1**	**0.1916**	**0.0667**	**0.0588**	**0.3219**	**8.26**	**0.0040**

Interaction is statistically significant "additive interaction".
Reject H0: IC = 0, i.e. Reject parallel lines on probability scale

Testing for additive interaction on the probability scale
Strategy #1b: Use linear normal (i.e. OLS regression) with robust standard errors.

****Weighted least squares – controls for the fact that not all observations have the same error variance using the Huber white heteroskedastic error estimation;

proc reg data = individual;

model depressn = stressevent vulnerability interaction/ white;

The REG Procedure
Model: MODEL1
Dependent Variable: depressn

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	--Heteroscedasticity Consistent- Standard Error	t Value	Pr > |t|
Intercept	1	0.01036	0.01884	0.55	0.5825	0.00729	1.42	0.1559
stressevent	1	0.09191	0.03366	2.73	0.0066	0.03311	2.78	0.0058
vulnerability	1	0.02190	0.03820	0.57	0.5668	0.02359	0.93	0.3539
interaction	**1**	**0.19162**	**0.05602**	**3.42**	**0.0007**	**0.06666**	**2.87**	**0.0043**

Interaction is statistically significant "additive interaction".
Reject H0: IC = 0, i.e. Reject parallel lines on probability scale

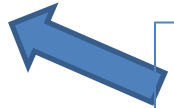
Strategy #2 Using logistic regression and back-transform to probability scale

Step 1 - Fit a logistic regression model with a cross product included

Step 2 – Back-transform to get predicted probabilites and then form IC contrast and do test of IC = 0.

Two ways of Back-transforming in the presence of covariates:

1) marginal predicted probabilities – get predicted probability at all covariate values and average across them
2) Conditional predicted probabilities – get predicted probability at fixed value of covariates (e.g. means or mode)



Greenland 2004 argues for marginal

Strategy #2 Using logistic regression and back-transform to probability scale

PROS:

- Probabilities are kept between 0 and 1.
- Can be done easily in STATA and SUDAAN (but not SAS)

CONS:

- **back-transforming can be tricky** for estimator and standard errors particularly in presence of covariates
- Homogeneity of covariate effects on odds ratio scale is not the same as homogeneity on risk difference scale and this may imply misspecification (Kalilani and Atashili 2006; Skrondal 2003)
- Not clear how to backtransform if either predictor is continuous.

Strategy #2 Using logistic regression and back-transform to probability scale

IN STATA

```
binreg depressn i.stress i.vul i.stress#i.vul, or
margins i.stress i.vul i.stress#i.vul, contrast(effects)
```

```
------------------------------------------------------------------------------
                           |            Delta-method
                           |   Contrast   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------------+--------------------------------------------------------------
                    stress |
              (1 vs base)  |   .1550216   .0292602     5.30   0.000     .0976726    .2123706
                           |
             vulnerability |
              (1 vs base)  |   .0968975   .0283118     3.42   0.001     .0414075    .1523876
                           |
      stress#vulnerability |
(1 vs base) (1 vs base)    |   .1916214   .0666555     2.87   0.004      .060979    .3222637
------------------------------------------------------------------------------
```

Interaction is statistically significant “additive interaction”.
Reject H0: IC = 0, i.e. Reject parallel lines on probability scale

RELATIONSHIP BETWEEN ADDITIVE AND MULTIPLICATIVE INTERACTION

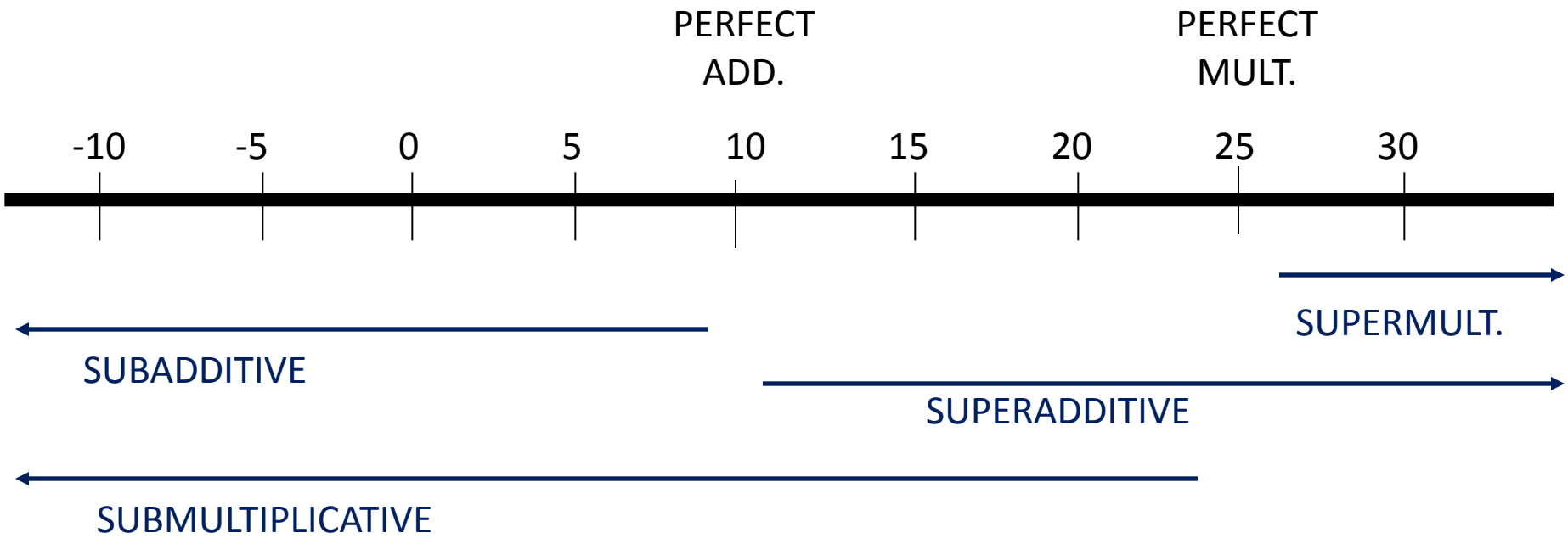
THEY CAN GIVE VERY DIFFERENT ANSWERS

No Interaction
Perfect Additivity

Risk Increment A = 5
Risk Increment B = 5
Risk Increment Both = 10

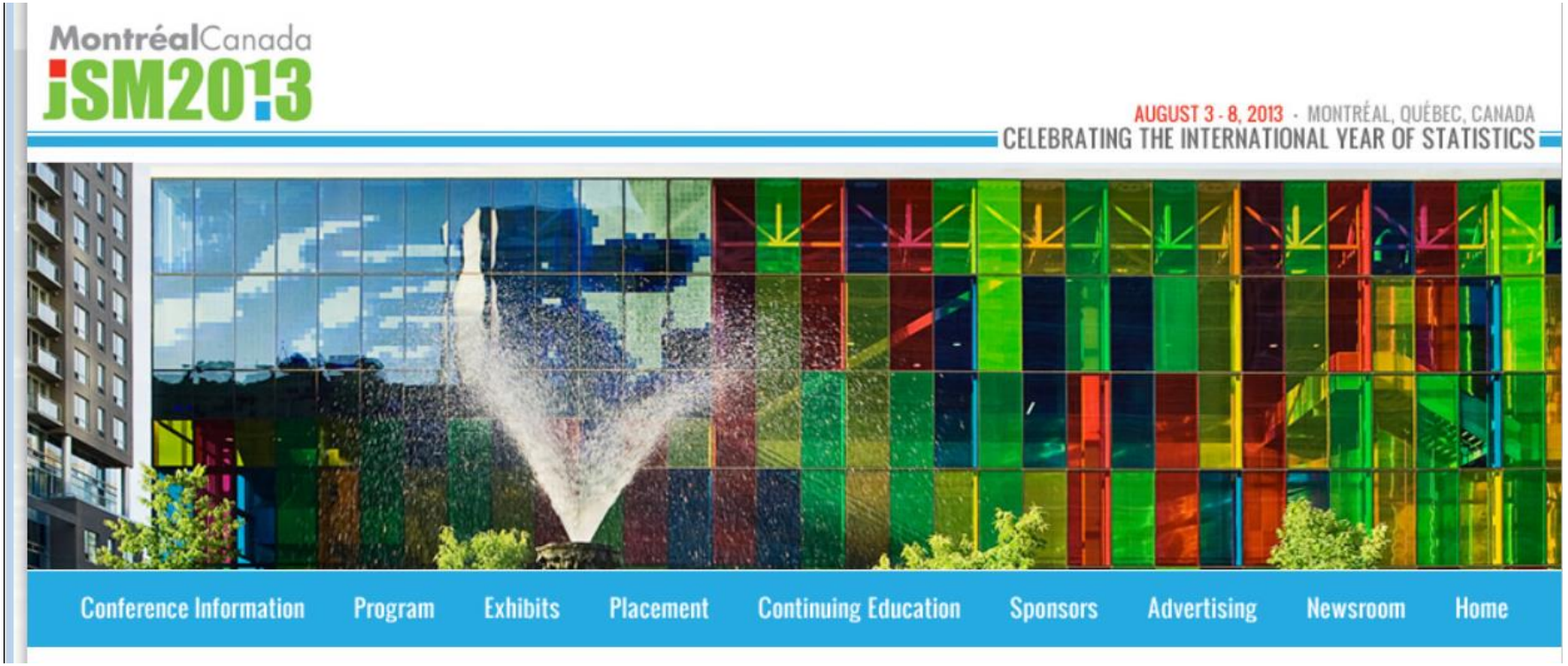
No interaction
Perfect Multiplicativity

Risk Increment A = 5
Risk Increment B = 5
Risk Increment Both = 25



Conclusion

- The appropriate scale on which to assess interaction effects with dichotomous outcomes has been a controversial topic in epidemiology for years, but **awareness of this controversy is not yet wide spread enough**.
- This would not be a problem if the status quo for examining effect modification (i.e. testing interaction effects in logistic regression) was actually the "RIGHT" thing to do, but, persuasive arguments have been made from the sufficient cause framework that the additive probability scale (not the multiplicative odds ratio scale) should be used to assess the presence of synergistic effects (Darroch 1997, Rothman and Greenland 1998, Schwartz 2006, Vanderwheel and Robins 2007,2008)
- There are now straightforward ways within existing software to estimate and test the statistical significance of additive interaction effects.
- **Additional work is needed getting the word out that effect modification should not (just) be looked at using Odds Ratios.**



Thank you for your attention.

See you in Montréal

Are you looking for the right interactions?

Statistically testing for interaction effects with dichotomous outcome variables

Updated 2-14-2012 for presentation to the Epi Methods group at Columbia

Melanie M. Wall
Departments of Psychiatry and Biostatistics
New York State Psychiatric Institute and Mailman School of Public Health
Columbia University
mmwall@columbia.edu

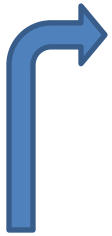
Data from Brown and Harris (1978) – 2X2X2 Table

Vulnerability	Exposure	Outcome (Depression)			
Lack of Intimacy	Stress Event	No	Yes	Risk of Depression	
No	No	191	2	0.010	P00
	Yes	79	9	0.102	P01
Effect of Stress given No Vulnerability ->		OR = 10.9 (2.3, 51.5) RR = 9.9 (2.2, 44.7)		RD = 0.092 (0.027,0.157)	
Yes	No	60	2	0.032	P10
	Yes	52	24	0.316	P11
Effect of Stress given Vulnerability ->		OR = 13.8 (3.1,61.4) RR = 9.8 (2.4, 39.8)		RD=0.284 (0.170,0.397)	

OR = Odds Ratio (95% Confidence Interval) <-compare to 1
RR = Risk Ratio (95% Confidence Interval) <-compare to 1
RD = Risk Difference (95% Confidence Interval) <-compare to 0

Does Vulnerability Modify the Effect of Stress on Depression?

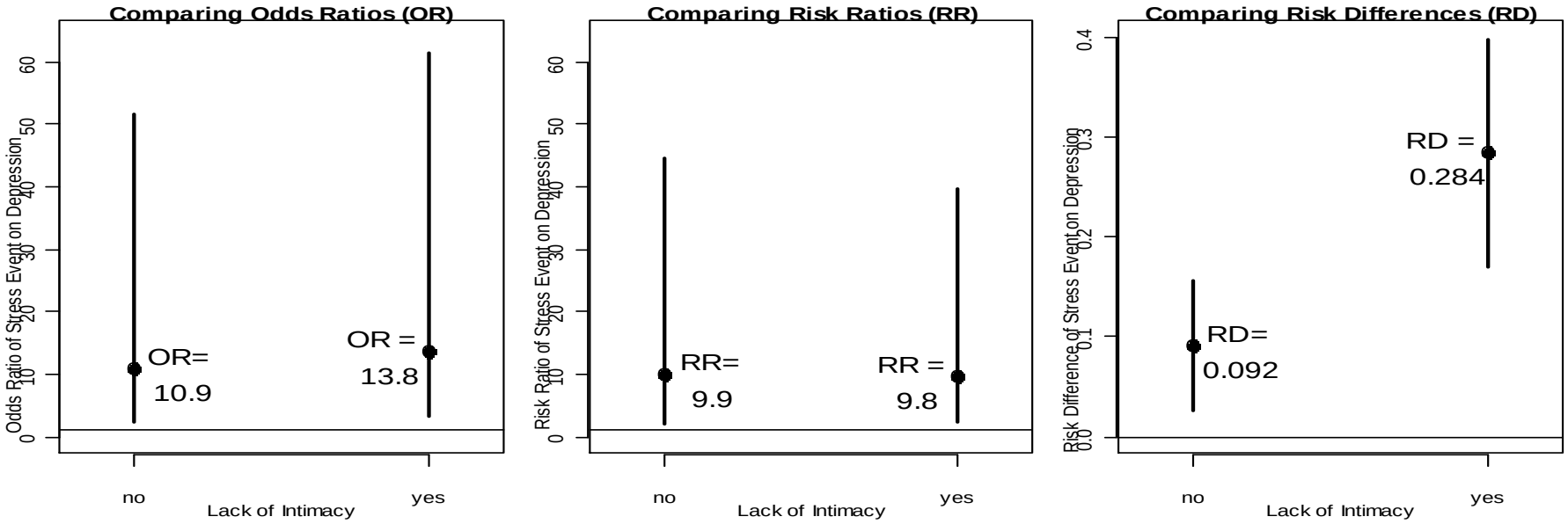
- On the **multiplicative Odds Ratio scale**, is 10.9 sig different from 13.8?
  - Test whether the ratio of the odds ratios (i.e. 13.8/10.9 = 1.27) is significantly different from 1.
- On the **multiplicative Risk Ratio scale**, is 9.9 sig different from 9.8?
  - Test whether the ratio of the risk ratios (i.e. 9.8/9.9 = 0.99) is significantly different from 1.
- On the **additive Risk Difference scale**, is 0.092 sig different from 0.284?
  - Test whether the difference in the risk differences (i.e. 0.28-0.09 = 0.19) is significantly different from 0.



Rothman calls this difference in the risk differences the "interaction contrast (IC)"

IC = (P11 - P10) – (P01 - P00)

Comparing stress effects across vulnerability groups
Different conclusions on multiplicative vs additive scale



95% confidence intervals for Odds Ratios overlap

-> **no statistically significant multiplicative interaction OR scale**

95% confidence intervals for Risk Ratios overlap

-> **no statistically significant multiplicative interaction RR scale**

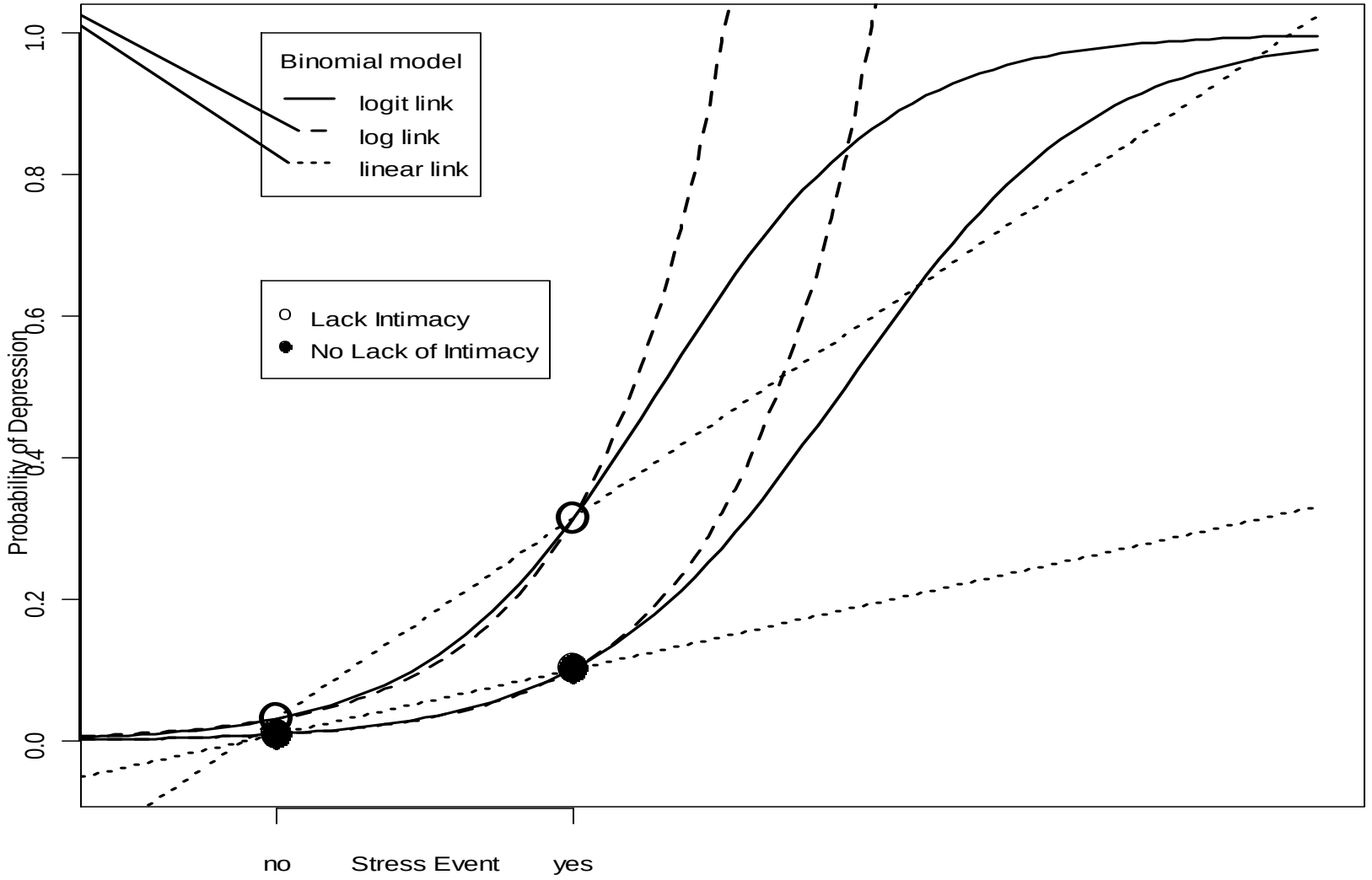
95% confidence intervals for Risk Differences do not overlap

-> **statistically significant additive interaction**

In general, it is possible to reach different conclusions on the two different multiplicative scales "distributional interaction" (Campbell, Gatto, Schwartz 2005)

Modeling Probabilities

Binomial modeling with logit, log, or linear link



Test for multiplicative interaction on the OR scale- Logistic Regression with a cross-product

IN SAS:

```
proc logistic data = brownharris descending;
model depressn = stressevent   lack_intimacy   stressevent*lack_intimacy;
oddsratio stressevent / at(lack_intimacy = 0 1);
oddsratio lack_intimacy / at(stressevent = 0 1);
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.5591	0.7108	41.1409	<.0001
stressevent	1	2.3869	0.7931	9.0576	0.0026
lack_intimacy	1	1.1579	1.0109	1.3120	0.2520
stresseve*lack_intim	**1**	**0.2411**	**1.0984**	**0.0482**	**0.8262**

exp(.2411) = 1.27 =
Ratio of Odds ratios =13.846/10.880
Not significantly different from 1

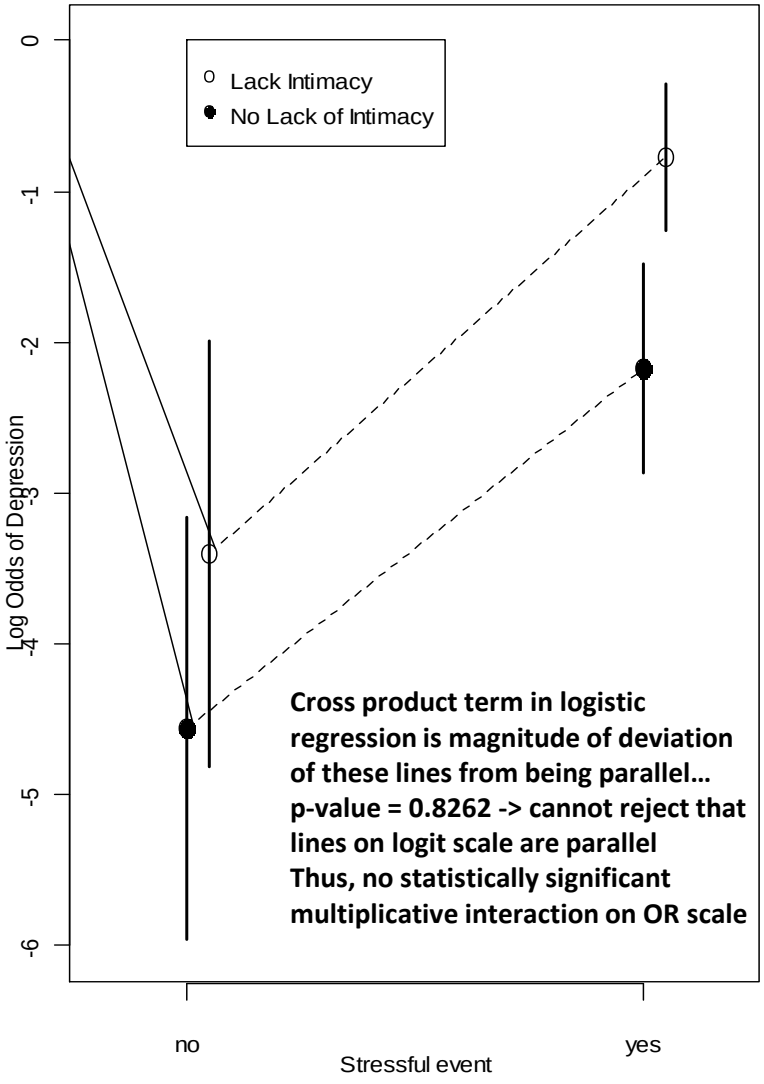
Wald Confidence Interval for Odds Ratios

Label	Estimate	95% Confidence Limits	
stressevent at lack_intimacy=0	10.880	2.299	51.486
stressevent at lack_intimacy=1	13.846	3.122	61.408
lack_intimacy at stressevent=0	3.183	0.439	23.086
lack_intimacy at stressevent=1	4.051	1.745	9.405

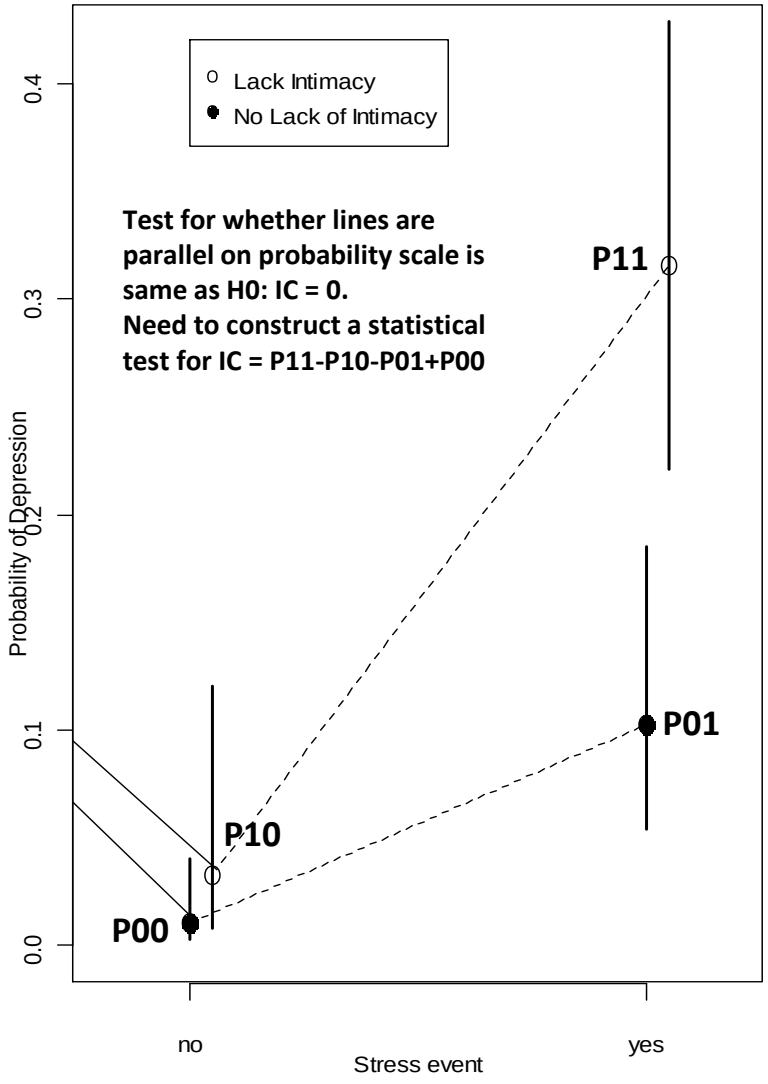
"multiplicative interaction" on OR scale is not significant

Test for interaction: Are the lines Parallel?

Log Odds scale



Probability scale



The Problem with Comparing Statistical Significance of Effects Across Groups

- Don't fall into the trap of concluding there must be effect modification because one association was statistically significant while the other one was not.
- In other words, just because a significant effect is found in one group and not in the other, does NOT mean the effects are necessarily different in the two groups (regardless of whether we use OR, RR, or RD).
- Remember, statistical significance is not only a function of the effect (OR, RR, or RD) but also the sample size and the baseline risk. Both of these can differ across groups.
- McKee and Vilhjalmsson (1986) point out that Brown and Harris (1978) wrongfully applied this logic to conclude there was statistical evidence of effect modification (fortunately there conclusion was correct despite an incorrect statistical test)

Different strategies for statistically testing additive interactions on the probability scale

The IC is the Difference of Risk Differences. IC = (P11 - P10) – (P01 - P00) = P11-P10-P01+P00

From Cheung (2007) "Now that many commercially available statistical packages have the capacity to fit log binomial and linear binomial regression models, 'there is no longer any good justification for fitting logistic regression models and estimating odds ratios' when the odds ratio is not of scientific interest" Inside quote from Spiegelman and Herzmark (2005).

1. Directly fit Risk = b0 + b1 * EXPO + b2 * VULN + b3*EXPO*VULN using (A) **linear binomial** or (B) **linear normal** model (but use robust standard errors). The b3 = IC and so a test for coefficient b3 is a test for IC. Can be implemented directly in PROC GENMOD or PROC REG. **PROS:** Contrast of interest is directly estimated and tested and covariates easily included **CONS:** Linear model for probabilities can be greater than 1 and less than 0 and thus maximum likelihood estimation can be a problem. Note there is no similar problem of estimation for the linear normal model. Wald-type confidence intervals can have poor coverage for linear binomial (Storer et al 1983), better to use profile likelihood confidence intervals.

2. Fit a **logistic regression** log(Risk/(1-Risk)) = b0 + b1 * EXPO + b2 * VULN + b3*EXPO*VULN, then **back-transform** parameters to the probability scale to calculate IC. Can be implemented directly in PROC NLMIXED. **PROS:** logistic model more computationally stable since smooth decrease/increase to 0 and 1. **CONS**: **back-transforming can be tricky** for estimator and standard errors particularly in presence of covariates. Covariate adjusted probabilities are obtained from average marginal predictions in the fitted logistic regression model (Greenland 2004). Homogeneity of covariate effects on odds ratio scale is not the same as homogeneity on risk difference scale and this may imply misspecification (Kalilani and Atashili 2006; Skrondal 2003).

3. Instead of IC, use IC ratio. Divide the IC by P00 and get a contrast of risk ratios:

   IC Ratio = P11/P00 -P10/P00 -P01/P00+P00/P00 = RR(11) – RR(10) – RR(01) + 1 called the

   Relative Excess Risk due to Interaction (RERI). **←Many papers on inference for RERI**

Testing for additive interaction on the probability scale
Strategy #1a: Use linear binomial regression with a cross-product

Risk = b0 + b1 * STRESS + b2 * LACKINT + b3*STRESS*LACKINT NOTE: b3 = IC

link=identity dist=binomial tells SAS to do linear binomial regression. lrci outputs likelihood ratio (profile likelihood) confidence intervals.

```
IN SAS:
proc genmod data = individual descending;
model depressn = stressevent   lack_intimacy   stressevent*lack_intimacy/
link = identity dist = binomial lrci;
estimate 'RD of stressevent when intimacy = 0' stressevent 1;
estimate 'RD of stressevent when intimacy = 1' stressevent 1 stressevent*lack_intimacy 1;
run;
```

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Likelihood Ratio 95% Confidence Limits		Wald Chi-Square	Pr>ChiSq
Intercept	1	0.0104	0.0073	0.0017	0.0317	2.02	0.1551
stressevent	1	0.0919	0.0331	0.0368	0.1675	7.70	0.0055
lack_intimacy	1	0.0219	0.0236	-0.0139	0.0870	0.86	0.3534
stresseve*lack_intim	**1**	**0.1916**	**0.0667**	**0.0588**	**0.3219**	**8.26**	**0.0040**

Contrast Estimate Results

Label	Mean Estimate	Mean Confidence Limits		Standard Error
RD of stressevent when intimacy = 0	0.0919	0.0270	0.1568	0.0331
RD of stressevent when intimacy = 1	0.2835	0.1701	0.3969	0.0578

Interaction is statistically significant "additive interaction".

Reject H0: IC = 0, i.e. Reject parallel lines on probability scale

Testing for additive interaction on the probability scale
Strategy #1b: Use linear normal (i.e. OLS regression) with robust standard errors.

****Weighted least squares – controls for the fact that not all observations have the same error variance;

proc reg data = individual;

model depressn = stressevent lack_intimacy interaction/ white; ***white does the huber white heteroskedastic error estimation;

run;

The REG Procedure
Model: MODEL1
Dependent Variable: depressn

Parameter Estimates

						--Heteroscedasticity Consistent-		
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standard Error	t Value	Pr > |t|
Intercept	1	0.01036	0.01884	0.55	0.5825	0.00729	1.42	0.1559
stressevent	1	0.09191	0.03366	2.73	0.0066	0.03311	2.78	0.0058
lack_intimacy	1	0.02190	0.03820	0.57	0.5668	0.02359	0.93	0.3539
interaction	1	0.19162	0.05602	3.42	0.0007	0.06666	2.87	0.0043

Test for additive interaction on the probability scale
Strategy #2: Use logistic regression and back-transform estimates to form contrasts on the probability scale

```
PROC NLMIXED DATA=individual;
***logistic regression model is;
odds = exp(b0 +b1*stressevent + b2*lack_intimacy + b3*stressevent*lack_intimacy);
pi = odds/(1+odds);
MODEL depressn~BINARY(pi);

estimate 'p00' exp(b0)/(1+exp(b0));
estimate 'p01' exp(b0+b1)/(1+exp(b0+b1));
estimate 'p10' exp(b0+b2)/(1+exp(b0+b2));
estimate 'p11' exp(b0+b1+b2+b3)/(1+exp(b0+b1+b2+b3));

estimate 'p11-p10' exp(b0+b1+b2+b3)/(1+exp(b0+b1+ b2+b3))-  exp(b0+b2)/(1+exp(b0+b2));
estimate 'p01-p00' exp(b0+b1)/(1+exp(b0+b1)) - exp(b0)/(1+exp(b0));

estimate 'IC= interaction contrast = p11-p10 - p01 + p00'
exp(b0+b1+b2+b3)/(1+exp(b0+b1+ b2+b3)) -  exp(b0+b2)/(1+exp(b0+b2)) - exp(b0+b1)/(1+exp(b0+b1)) +
exp(b0)/(1+exp(b0));

estimate 'ICR= RERI using RR = p11/p00 - p10/p00 - p01/p00 + 1'
        exp(b0+b1+b2+b3)/(1+exp(b0+b1+ b2+b3))/ (exp(b0)/(1+exp(b0)))
      -  exp(b0+b2)/(1+exp(b0+b2))/  (exp(b0)/(1+exp(b0)))
      -  exp(b0+b1)/(1+exp(b0+b1)) /  (exp(b0)/(1+exp(b0)))  + 1;
estimate 'ICR= RERI using OR'  exp(b1+b2+b3) - exp(b1) - exp(b2) +1;
RUN;
```

These are Strategy #3

Strategy #2 Output from NLMIXED

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
b0	-4.5591	0.7108	419	-6.41	<.0001	0.05	-5.9563	-3.1620	-0.00002
b1	2.3869	0.7931	419	3.01	0.0028	0.05	0.8280	3.9458	-0.00003
b2	1.1579	1.0109	419	1.15	0.2527	0.05	-0.8291	3.1450	2.705E-6
b3	0.2411	1.0984	419	0.22	0.8264	0.05	-1.9180	2.4002	-0.00001

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper
p00	0.01036	0.00728	419	1.42	0.1559	-0.00397	0.0246
p10	0.1023	0.03230	419	3.17	0.0017	0.03878	0.1658
p01	0.03226	0.02244	419	1.44	0.1513	-0.01185	0.0763
p11	0.3158	0.05332	419	5.92	<.0001	0.2110	0.4206
p11-p10	0.2135	0.06234	419	3.43	0.0007	0.09098	0.3361
p01-p00	0.02190	0.02359	419	0.93	0.3539	-0.02448	0.0682
IC =p11-p10-p01+p00	**0.1916**	**0.06666**	**419**	**2.87**	**0.0042**	**0.06060**	**0.3226**
RERI using RR	**18.4915**	**13.8661**	**419**	**1.33**	**0.1831**	**-8.7644**	**45.7473**
RERI using OR	**31.0138**	**24.3583**	**419**	**1.27**	**0.2036**	**-16.8659**	**78.8936**

IC estimator same as strategy #1, but slightly different s.e., p-value, 95% conf interval

- The IC estimator is same as before (slide 9) but slightly different s.e., p-value and 95% confidence interval – still **conclude there is a significant additive interaction**.

- **Results for RERI (using RR and OR) indicate that there is NOT a significant additive interaction**. This conflicts with the conclusion that the IC is highly significant. The cause of the discrepancy is related to estimation of standard errors and confidence intervals. Literature indicates Wald-type confidence intervals perform poorly for RERI (Hosmer and Lemeshow 1992; Assman et al 1996).

- Proc NLMIXED uses Delta method to obtain standard errors of back-transformed parameters and Wald-type confidence intervals, i.e. (estimate) +- 1.96*(standard error) . Possible to obtain profile likelihood confidence intervals using a separate macro (Richardson and Kaufman 2009) or PROC NLP (nonlinear programming) (Kuss et al 2010). Also possible to bootstrap (Assman et al 1996 and Nie et al 2010) or incorporate prior information (Chu et al 2011)

Test for additive interaction on the probability scale
Strategy #2: Use logistic regression and back-transform – An easier way in SUDAAN

```
proc rlogist data = a design = srs; ***srs tells SUDAAN to treat as iid data;
class gene expo ;
reflevel gene=0 expo=0;
model outcome= gene expo gene*expo;

**********;
predMARG gene*expo;
pred_eff gene=(1 0)*expo=(-1 1)/name ="pred: exposure effect when gene not present";
pred_eff gene=(0 1)*expo=(-1 1)/name ="pred: exposure effect when gene is present";
pred_eff gene=(-1 1)*expo=(-1 1)/name ="pred_int: difference in risk differences";

**********;
condMARG gene*expo;
cond_eff gene=(1 0)*expo=(-1 1)/name ="cond: exposure effect when gene not present";
cond_eff gene=(0 1)*expo=(-1 1)/name ="cond: exposure effect when gene is present";
cond_eff gene=(-1 1)*expo=(-1 1)/name ="cond_int: difference in risk differences";

run;
```

These are the same if there are no covariates

```
S U D A A N
            Software for the Statistical Analysis of Correlated Data
           Copyright      Research Triangle Institute     August 2008
                                 Release 10.0
DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a
Simple Random Sample (SRS) Design

Number of zero responses     :   382
Number of non-zero responses :    37
```

Response variable OUTCOME: OUTCOME
by: Predicted Marginal #1.

NOTE: I renamed
Gene = lack_intimacy
Expo = stress_event
But data is same as
Brown Harris

PredMarginal #1	Predicted Marginal	SE	T:Marg=0	P-value
GENE, EXPO				
0, 0	0.0103626943	0.0072981803	1.4199011081	0.1563818430
0, 1	0.1022727273	0.0323392291	3.1624973812	0.0016782922
1, 0	0.0322580645	0.0224658037	1.4358740502	0.1517860946
1, 1	0.3157894737	0.0533833440	5.9155056648	0.0000000069

Contrasted

Pred Marg#1	PREDMARG Contrast	SE	T-Stat	P-value
pred: exposure effect when gene not present	0.0919100330	0.0331525138	2.7723397823	0.0058143088

More SUDAAN output for the BrownHarris example

```
------------------------------------------------------------------------------
Contrasted
  Predicted
  Marginal     PREDMARG
  #2           Contrast                    SE           T-Stat          P-value
------------------------------------------------------------------------------
pred:
  exposure
  effect
  when gene
  is present     0.2835314092     0.0579179916     4.8953943573     0.0000014039
------------------------------------------------------------------------------

------------------------------------------------------------------------------
Contrasted
  Predicted
  Marginal     PREDMARG
  #3           Contrast                    SE           T-Stat          P-value
------------------------------------------------------------------------------
pred_int:
  difference
  in risk
  differenc-
  es             0.1916213762     0.0667351701     2.8713701643     0.0042948171
------------------------------------------------------------------------
```

Controlling for Covariates within the back-transformation strategy

- When we fit a model on the logit scale and include a covariate (e.g. age as a main effect)

  logit p = b0 + b1 * EXPO + b2 * VULN + b3*EXPO*VULN + b4*Gender

  this is controlling for gender on the logit scale so that the effects we find for expo, vuln, and expo*vuln on the logit scale (odds ratios) are expected to be the same both genders.
- BUT, back on the probability scale (after back-transformation), the effect of expo, vuln, or expo*vuln will differ depending on which gender someone is
- That is, there was “homogeneity” of effects (across gender) on the logit scale, but there will not be “homogeneity” of effects on the probability scale.
- Not sure this if this is a problem per se but it is something necessary to consider for interpretation.

Two ways to back transform – SUDAAN calls them "conditional and predicted" marginal

From p.513 SUDAAN manual

Conditional and Predicted Marginals (Model-Adjusted Risks)

The estimated ***conditional marginal*** from the logistic model is:

Back-transform an "average person"

$$\mathbf{M}_C = \frac{\exp(\bar{\mathbf{x}}'^{*}\hat{\boldsymbol{\beta}})}{1+\exp(\bar{\mathbf{x}}'^{*}\hat{\boldsymbol{\beta}})}, \text{ where } \bar{\mathbf{x}}^{*'} = \frac{\sum_{h=1}^{H}\sum_{i=1}^{I_h}\sum_{j=1}^{J_{hi}}\sum_{k=1}^{K_{hij}} w_{hijk}\mathbf{x}'^{*}_{\mathbf{hijk}}}{\sum_{h=1}^{H}\sum_{i=1}^{I_h}\sum_{j=1}^{J_{hi}}\sum_{k=1}^{K_{hij}} w_{hijk}},$$

Bieler (2010) Standardized population averaged risk Greenland (2004)

and the estimated ***predicted marginal*** is:

$$\mathbf{M}_P = \frac{1}{w_{++++}}\sum_{h=1}^{H}\sum_{i=1}^{I_h}\sum_{j=1}^{J_{hi}}\sum_{k=1}^{K_{hij}} w_{hijk}\,\frac{\exp(\mathbf{x}'^{*}_{\mathbf{hijk}}\,\hat{\boldsymbol{\beta}})}{1+\exp(\mathbf{x}'^{*}_{\mathbf{hijk}}\,\hat{\boldsymbol{\beta}})}.$$

Two ways of back-transforming…

- The conditional marginal approach estimates the effects on the probability scale for a certain fixed covariate value, usually the mean, but usually an "average person" doesn't exist, e.g. if the covariate is gender and if 30% of sample is male, this means we are finding the predicted probability associated with someone who is 30% male. Further, if we used some other fixed values for the covariates, we would get different effects (on the probability scale) for the gene*environment
- The predicted marginal approach allows different covariate values to give different predicted probabilities and thus gets a distribution of risks and then averages over them on the probability scale. SUDAAN uses the observed distribution of the covariates in the sample as the "standardization population", but I believe Sharon Schwartz would argue that perhaps it is better to use the distribution of the covariates in the "unexposed" population to standardize to. I think she had a student work on this for topic and they are working on a paper now.
- From my reading of the EPI literature, the predicted marginal approach is preferred since it has a more meaningful interpretation.

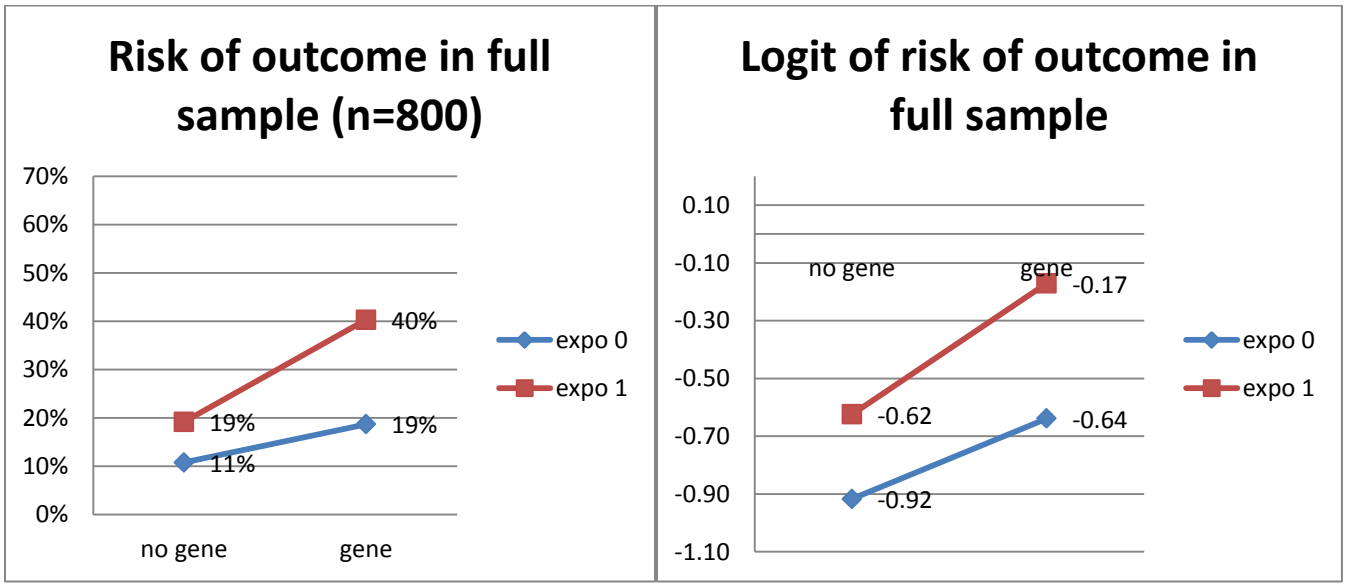
Conclusion

- The appropriate scale on which to assess interaction effects with dichotomous outcomes has been a controversial topic in epidemiology for years, but **awareness of this controversy is not yet wide spread enough**.
- This would not be a problem if the status quo for examining effect modification (i.e. testing interaction effects in logistic regression) was actually the "RIGHT" thing to do, but, persuasive arguments have been made from the sufficient cause framework that the additive probability scale (not the multiplicative odds ratio scale) should be used to assess the presence of synergistic effects (Darroch 1997, Rothman and Greenland 1998, Schwartz 2006, Vanderwheel and Robins 2007,2008)
- There are now straightforward ways within existing software to estimate and test the statistical significance of additive interaction effects.
- **Additional work is needed getting the word out that effect modification should not (just) be looked at using Odds Ratios.**

An illustrative Data example – Additive interaction with a covariate

gene	expo	N	% outcome	% of covariate = 1
0	0	241	11%	24%
0	1	172	19%	36%
1	0	171	19%	32%
1	1	216	40%	37%

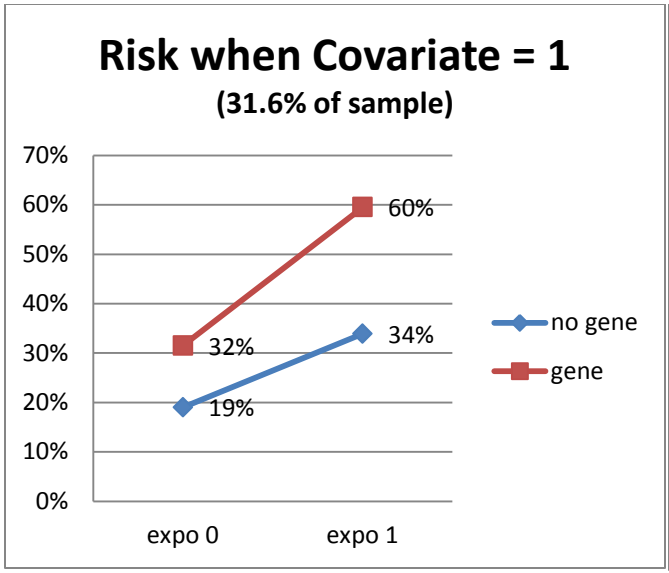
If there was perfect balance of the covariate in each of the 4 risk groups (i.e. Gene by exposure groups), then we wouldn't have to worry about it being a confounder, but since it varies from 24% up to 37% and tends to be higher in the exposure group as compared to the not exposed group, it should be controlled.

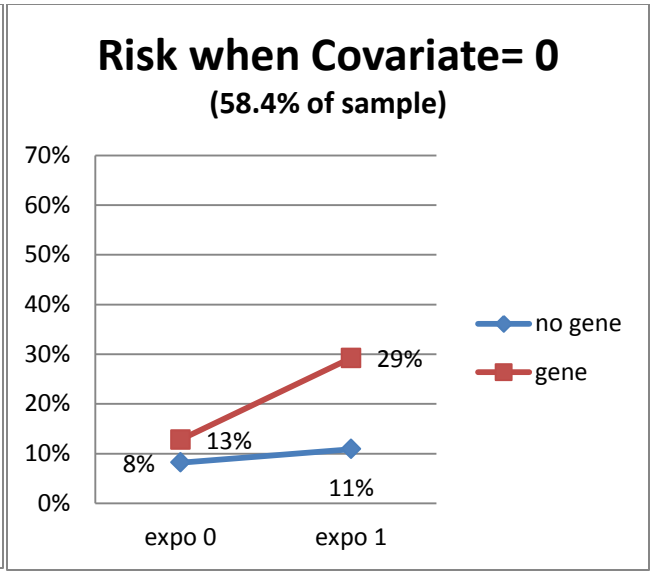


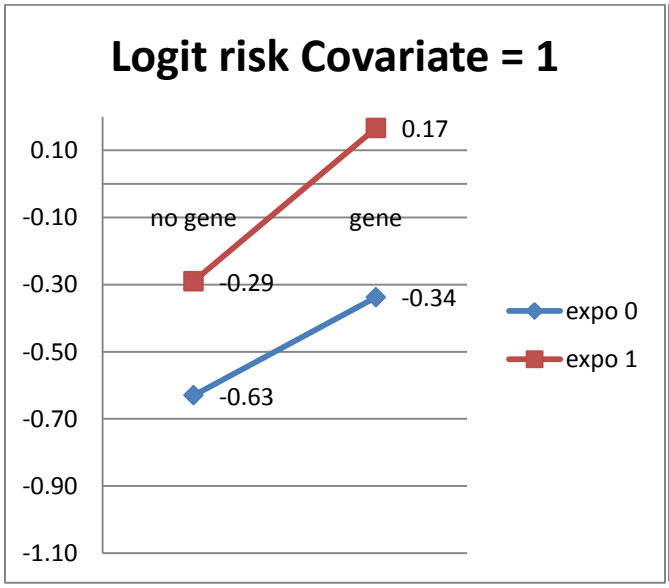
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
outcome	800	0.22250	0.41619	178.00000	0	1.00000
gene	800	0.48375	0.50005	387.00000	0	1.00000
expo	800	0.48500	0.50009	388.00000	0	1.00000
covariate	800	0.31625	0.46530	253.00000	0	1.00000

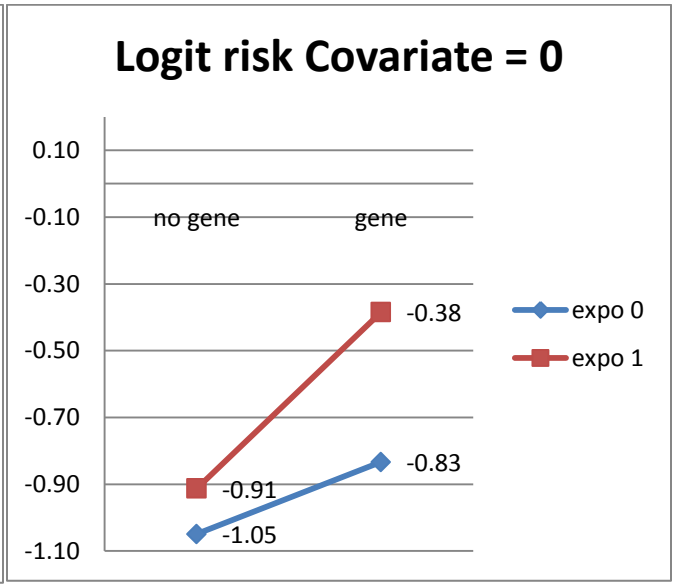
Pearson Correlation Coefficients, N = 800

	outcome	gene	expo	covariate
outcome	1.00000	0.19781	0.20247	0.25663
		<.0001	<.0001	<.0001
gene	0.19781	1.00000	0.14166	0.05708
	<.0001		<.0001	0.1067
expo	0.20247	0.14166	1.00000	0.09840
	<.0001	<.0001		0.0053
covariate	0.25663	0.05708	0.09840	1.00000
	<.0001	0.1067	0.0053	









Logistic regression estimates:

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Likelihood Ratio 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.5218	0.2264	-2.9885	-2.0983	124.09	<.0001
gene	1	0.5668	0.2926	-0.0050	1.1462	3.75	0.0527
expo	1	0.5383	0.2913	-0.0309	1.1155	3.41	0.0647
gene*expo	1	0.5514	0.3839	-0.2001	1.3075	2.06	0.1509
covariate	1	1.2192	0.1846	0.8589	1.5835	43.60	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		

Logit P = -2.52 + 0.5668*GENE + 0.5383*EXPO + 0.5514*G*E + 1.219*Covariate

The conditional marginal approach then calculates the predicted probabilities for the G*E back on the original scale as

```
> gene=c(0,0,1,1)
> expo=c(0,1,0,1)
> covariate = .31625
> LogitP = -2.52 + 0.5668*gene + 0.5383*expo + 0.5514*gene*expo + 1.219*covariate
> LogitP
[1] -2.1344912 -1.5961912 -1.5676912 -0.4779913

> exp(LogitP)/(1+exp(LogitP))
[1] 0.1057894 0.1685146 0.1725458 0.3827266
```

FROM SUDAAN

Conditional Marginal #1	Conditional Marginal	SE	T:Marg=0	P-value
GENE, EXPO				
0, 0	0.1056206853	0.0203578588	5.1882020754	0.0000002694
0, 1	0.1682595536	0.0280578276	5.9968845713	0.0000000030
1, 0	0.1722933591	0.0292915541	5.8820149491	0.0000000060
1, 1	0.3823099821	0.0342856575	11.1507262872	0.0000000000

IC = (.38 - .17) – (.168 - .105)

Contrasted Condition-al Marginal #3	CONDMARG Contrast	SE	T-Stat	P-value
cond_int: difference in risk differenc-es	0.1473777547	0.0563674058	2.6145917595	0.0091017247

And for the predicted marginal approach we take

```
> gene=c(0,0,1,1,0,0,1,1)
> expo=c(0,1,0,1,0,1,0,1)
> covariate = c(0,0,0,0,1,1,1,1)
> #covariate = .31625
> LogitP = -2.52 + 0.5668*gene + 0.5383*expo + 0.5514*gene*expo + 1.219*covariate
> LogitP
[1] -2.5200 -1.9817 -1.9532 -0.8635 -1.3010 -0.7627 -0.7342  0.3555
> prob=exp(LogitP)/(1+exp(LogitP))
> prob
[1] 0.07446795 0.12113773 0.12420485 0.29660861 0.21399677 0.31806035 0.32427374
[8] 0.58795068
> predmarg = (1-.31625)*prob[1:4] + .31625*prob[5:8]
> predmarg
[1] 0.1185939 0.1834145 0.1874766 0.3887455
```

Predicted Marginal #1	Predicted Marginal	SE	T:Marg=0	P-value
GENE, EXPO				
0, 0	0.1184185808	0.0214723290	5.5149388193	0.0000000471
0, 1	0.1831610712	0.0281468459	6.5073391217	0.0000000001
1, 0	0.1872265641	0.0292713633	6.3962365483	0.0000000003
1, 1	0.3883600906	0.0320980597	12.0991765126	0.0000000000

IC = .388 - .187 – (.183 - .118) = .136

Contrasted Predicted Marginal #3	PREDMARG Contrast	SE	T-Stat	P-value
pred_int: difference in risk differenc-es	0.1363910361	0.0555247836	2.4563992369	0.0142454402