Effect measure modification & Interaction

Madhukar Pai, MD, PhD
McGill University
madhukar.pai@mcgill.ca
Interaction + Effect Modification = Frustration

EDITORIAL

Interaction: A word with two meanings creates confusion

Anders Ahlbom & Lars Alfredsson
Institute of Environmental Medicine and Stockholm Center for Public Health, Box 210 171 77, Stockholm, Sweden

Accepted in revised form 22 March 2005

Perhaps more than any other word in epidemiology, ‘interaction’ presents a challenge to clinical and epidemiological researchers. The problem stems from its applicability to describe two different phenomena. On the one hand, interaction refers to the biologic interaction of two or more causes of disease that together assert their influence on disease risk. On the other, interaction refers to statistical interaction which is the necessity for a product term in a linear model. In this editorial, we have two related goals: (1) dependent variables is no longer additive. A logistic regression model on the other hand is implicitly exponential and thus multiplicative. It becomes additive only after a logarithmic transformation. As a consequence, the inclusion of an interaction term in the logistic regression model implies that the investigated relation is no longer multiplicative. The confusion around the dual meaning of the term interaction has arisen in parallel with the widespread use of statistical modeling and software

“Introduction to effect modification leaves some students of epidemiology struggling with the distinction between this and the other 'third variable' phenomenon, namely, confounding. Confusion regarding effect modification is further exacerbated by a lack of consensus on both semantic and conceptual issues” Joseph KS. Paediatr Perinat Epidemiol. 2009

“The term “interaction” is a minefield of potential misunderstanding…the presentation and discussion of interaction in the medical and epidemiologic literature is woefully inadequate.” JS Kaufman, Epidemiol 2009
**Terminology**

- **Biological interaction**

- **Effect modification**
  - Or, more precisely, “effect-measure modification”

- **Heterogeneity of effects**

- **Subgroup effects** (i.e. effect varies across subgroups)

- **Statistical Interaction**
  - Deviation from a specified model form (additive or multiplicative)

Synonymous

Often used interchangeably
On the Distinction Between Interaction and Effect Modification

Tyler J. VanderWeele

Abstract: This paper contrasts the concepts of interaction and effect modification using a series of examples. Interaction and effect modification are formally defined within the counterfactual framework. Interaction is defined in terms of the effects of 2 interventions whereas effect modification is defined in terms of the effect of one intervention varying across strata of a second variable. Effect modification can be present with no interaction; interaction can be present with no effect modification. There are settings in which it is possible to assess effect modification but not interaction, or to assess interaction but not effect modification. The analytic procedures for obtaining estimates of effect modification parameters and interaction parameters using marginal structural models are compared and contrasted. A characterization is given of the settings in which interaction and effect modification coincide.

(Epidemiology 2009;20: 863–871)

of what will be formally defined below as an interaction of effects. Sometimes the coefficient for the product term can be interpreted both as a measure of effect modification and as a measure of interaction; sometimes only one of the 2 interpretations (or neither) is warranted.

The paper is structured as follows. First, I provide and contrast formal counterfactual definitions for interaction and effect modification. Second, examples are given showing that it is possible to have effect modification without interaction or interaction without effect modification. Third, further examples are given showing that in some cases it is possible to identify effect modification but not interaction and that in other cases it is possible to identify interaction but not effect modification. Fourth, analytic procedures to estimate interaction and effect modification parameters in marginal structural
Biological interaction

“the interdependent operation of two or more biological causes to produce, prevent or control an effect”

[Porta, Dictionary, 2008]
Multicausality and interdependent effects

- Disease processes tend to be multifactorial: “multicausality”
  - Very few exposures cause disease entirely by themselves
    - Exposure to measles can cause measles only if somebody is susceptible (e.g. not vaccinated)
    - Development of melanoma among those with high UV light exposure who also have fair skin
- The “one-variable-at-a-time” perspective has several limitations
- Both confounding and effect modification are manifestations of multicausality (reality is multivariate!)

Schoenbach, 2000
Biological interaction

- Refers to “co-participation in a causal mechanism of two or more component causes” (Rothman 2002)
- Illustrated by the “causal pie” model (Rothman)
Biological interaction

Figure 2–1. Three sufficient causes of a disease.
Example: Phenylketonuria

- PLU is a condition linked to a dietary factor (phenylalanine) and a genetic defect (mutations in the structural gene for phenylalanine hydroxylase)

Dietary factor

PKU gene mutation

Unknown factors
Drinking & Driving = Lethal Interaction!
Effect modification, statistical interaction, heterogeneity of effects
Effect modification & statistical interaction

- Two definitions (but related):
  - Definition based on homogeneity or heterogeneity of effects
    - Interaction occurs when the effect of a risk factor (X) on an outcome (Y) is not homogeneous in strata formed by a third variable (Z, effect modifier)
    - “Differences in the effect measure for one factor at different levels of another factor” [Porta, 2008]
    - This is often called “effect modification”
  - Definition based on the comparison between observed and expected joint effects of a risk factor and a third variable [deviation from some specified model]
    - Interaction occurs when the observed joint effects of the risk factor (X) and third variable (Z) differs from that expected on the basis of their independent effects
    - This is often called “statistical interaction”
Definition based on homogeneity or heterogeneity of effects

This is most commonly called “effect modification”
Definition based on homogeneity or heterogeneity of effects

- Effect of exposure on the disease is modified (altered) depending on the value of a third variable called “effect modifier”
Crude 2 x 2 table

Calculate Crude RR, OR

Stratify by 3rd variable

Calculate RR’s, OR’s for each stratum

Test whether stratum-specific RR’s, OR’s are similar (test for homogeneity)

If they are similar, investigate the possibility the 3rd variable is a confounder.

If they are different, there is evidence of effect modification.
Evaluation of confounding and interaction

Are stratum-specific RR’s similar?

- **YES**
  - crude RR = adjusted RR?
    - **NO**
      - CONFOUNDING
      - report adjusted measure (e.g. MH RR)
    - **YES**
      - NO CONFOUNDING
      - report crude OR or RR

- **NO**
  - INTERACTION… report stratum-specific OR or RR
Decision tree for evaluating confounding and effect modification

1. Calculate an unadjusted (collapsed) estimate of association between the exposure and outcome of interest (OR, CIR, IDR, RD)

2. Stratify the data based on the potential confounder or effect modifier

3. Calculate stratified estimates of association (OR, CIR, IDR, RD)

4. Compare the stratified estimates using the test for homogeneity

5. Are the estimates the same?
   
   Yes  No

6. Calculate an adjusted estimate using Mantel-Haenszel method  Effect Modification Present
   Report stratified estimates

7. Compare the adjusted estimate to the unadjusted estimate (from step 1)

8. Are the estimates the same?
   
   Yes  No

9. No confounding present  Confounding present
   Report unadjusted estimate  Report adjusted estimate
1. Calculate crude RR
2. Stratify and calculate Stratum-specific RR
   - Stratum-specific RR are similar
     - No effect modification
       - Calculate pooled RR
         - Crude RR \approx Adjusted RR
           - No major confounding
             - Use Crude RR
         - Crude RR \neq Adjusted RR
           - Confounding present
             - Use adjusted RR
   - Stratum-specific RR are different
     - Effect modification
       - Use stratum-specific RR
Two 'average' men having an 'average' meal.
Confounding versus interaction

- Confounding is a problem we want to eliminate (control or adjust for) in our study
  - Evaluated by comparing crude vs. adjusted effect estimates: is the adjusted estimate different from the crude one?

- Interaction is a natural occurrence that we want to describe and study further
  - Evaluated by comparing stratum-specific estimates: are they different from one another?
Example: Smoking and myocardial infarction (MI)

1) Calculate crude measure of association...

<table>
<thead>
<tr>
<th></th>
<th>MI</th>
<th>no MI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smokers</td>
<td>42</td>
<td>158</td>
<td>200</td>
</tr>
<tr>
<td>Nonsmokers</td>
<td>21</td>
<td>175</td>
<td>196</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>63</td>
<td>333</td>
<td>396</td>
</tr>
</tbody>
</table>

$OR = \frac{ad}{bc}$

$OR = 2.22 \ (95\%\ CI\ 1.26,\ 3.91)$

Investigators decided to look at dietary fat as a confounder/effect modifier.
2) Calculate stratum-specific measures of association...

**STRATUM 1: Dietary fat consumption <30% of calories**

<table>
<thead>
<tr>
<th></th>
<th>MI</th>
<th>noMI</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smokers</td>
<td>12</td>
<td>133</td>
<td>145</td>
</tr>
<tr>
<td>Nonsmokers</td>
<td>11</td>
<td>123</td>
<td>134</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>256</td>
<td>279</td>
</tr>
</tbody>
</table>

OR = 1.01
(0.429, 2.37)

**STRATUM 2: Dietary fat consumption > 30% of calories**

<table>
<thead>
<tr>
<th></th>
<th>MI</th>
<th>noMI</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smokers</td>
<td>30</td>
<td>25</td>
<td>55</td>
</tr>
<tr>
<td>Nonsmokers</td>
<td>10</td>
<td>52</td>
<td>62</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>77</td>
<td>117</td>
</tr>
</tbody>
</table>

OR = 6.29
(2.64, 14.75)
Inference

• CRUDE OR for smoking and MI = 2.22

• STRATUM-SPECIFIC OR for smoking and MI with dietary fat consumption as a potential interacting variable...
  DFC<30% OR = 1.01 (0.425, 2.37)
  DFC>30% OR = 6.29 (2.64, 14.75)

– Is there effect modification?
– Is there confounding?
– Which measure should we report?
More numeric examples

<table>
<thead>
<tr>
<th>Study</th>
<th>Crude RR</th>
<th>Stratum1 RR</th>
<th>Stratum2 RR</th>
<th>Interaction?</th>
<th>Confounding?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.00</td>
<td>1.02</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.00</td>
<td>1.02</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.70</td>
<td>0.03</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4.10</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.20</td>
<td>4.00</td>
<td>4.10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Real example: VaxGen HIV Vaccine Trial

<table>
<thead>
<tr>
<th>All subjects</th>
<th>Total</th>
<th>Infected at end of trial</th>
<th>Percentage who became infected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Placebo</td>
</tr>
<tr>
<td></td>
<td>1,679</td>
<td>98</td>
<td>5.8%</td>
</tr>
<tr>
<td></td>
<td>3,330</td>
<td>191</td>
<td>5.7%</td>
</tr>
<tr>
<td>White &amp; Hispanic</td>
<td></td>
<td></td>
<td>Vaccine</td>
</tr>
<tr>
<td></td>
<td>1,508</td>
<td>81</td>
<td>5.4%</td>
</tr>
<tr>
<td></td>
<td>3,003</td>
<td>179</td>
<td>6.0%</td>
</tr>
<tr>
<td>Black, Asian, other combined</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>171</td>
<td>17</td>
<td>9.9%</td>
</tr>
<tr>
<td></td>
<td>327</td>
<td>12</td>
<td>3.7%</td>
</tr>
<tr>
<td>Black</td>
<td>111</td>
<td>9</td>
<td>8.1%</td>
</tr>
<tr>
<td></td>
<td>203</td>
<td>4</td>
<td>2.0%</td>
</tr>
</tbody>
</table>

Source: VaxGen, Inc.
Example: VaxGen HIV Vaccine Trial

Risk Ratio for all participants: 5.7% / 5.8% = 0.98
No protection

Risk Ratio for African Americans: 2.0% / 8.1% = 0.25
75% protection!

Effect modification: race modifies the effect of HIV vaccine
“race” is the “effect measure modifier” (using RR as effect measure)
Other examples

- Age and measles vaccination
- Smoking during pregnancy, birth weight, and maternal age
- Smoking, oral contraceptives, and myocardial infarction
- Cardiovascular risks of HRT: years since menopause
- Race and antihypertensive medications
- Circumcision and HIV: heterosexual vs MSM
Comparison of Vaccination with Measles-Mumps-Rubella Vaccine at 9, 12, and 15 Months of Age

Stephen C. Redd,1 Gail E. King,1* Janet L. Heath,2 Baghar Forghani,2 William J. Bellini,3 and Lauri E. Markowitz1

1National Immunization Program and 2National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia; 3California State Department of Health Services, Viral and Rickettsial Disease Laboratory, Richmond, California

To determine seroconversion rates with measles-mumps-rubella vaccine administered to children at 9, 12, or 15 months of age, we undertook a prospective randomized trial. Among children vaccinated at 15 months of age, 98% seroconverted to measles, compared with 95% of those vaccinated at 12 months of age and 87% of those vaccinated at 9 months of age. In each age group, children of mothers born in or before 1963 had lower rates of seroconversion against measles, with the lowest rate in children vaccinated at 9 months. The seroconversion rate of rubella paralleled that of measles, with the lowest seroconversion rates in children vaccinated at 9 months of age whose mothers were born in or before 1963. The response to mumps varied little by age of the child or birth year of the child’s mother. These results support the recommended age for first vaccination with measles-mumps-rubella at 12–15 months.

Table 2. Seroconversion by vaccine antigen, age group vaccinated, and birth year of mother among children vaccinated with measles-mumps-rubella vaccine at 9, 12, or 15 months of age.

<table>
<thead>
<tr>
<th>Vaccine antigen, birth year of mother</th>
<th>Randomization age group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9 months</td>
</tr>
<tr>
<td>Measles, overall</td>
<td>249/285 (87.4%)</td>
</tr>
<tr>
<td>1963 or earlier</td>
<td>147/176 (83.5%)</td>
</tr>
<tr>
<td>After 1963</td>
<td>102/109 (93.6%)</td>
</tr>
<tr>
<td>Rubella, overall</td>
<td>249/273 (91.2%)</td>
</tr>
<tr>
<td>1963 or earlier</td>
<td>148/167 (88.6%)</td>
</tr>
<tr>
<td>After 1963</td>
<td>101/106 (95.3%)</td>
</tr>
<tr>
<td>Mumps, overall</td>
<td>251/272 (92.3%)</td>
</tr>
<tr>
<td>1963 or earlier</td>
<td>154/167 (92.2%)</td>
</tr>
<tr>
<td>After 1963</td>
<td>97/105 (92.4%)</td>
</tr>
</tbody>
</table>

**NOTE.** Data are no/total (%).

Age modifies the efficacy of MMR vaccination
Birth Weight and Smoking During Pregnancy—Effect Modification by Maternal Age

Steven H. Fox,1,2 Thomas D. Koepsell,1 and Janet R. Daling1

Cigarette smoking during pregnancy is an important, avoidable factor associated with low birth weight. Maternal age is also associated with variations in birth weight. Using birth certificate data from all 347,650 singleton births for which maternal age and birth weight were recorded during 1984–1988 in Washington State, this study investigated birth weight and smoking during pregnancy (yes/no) for mothers of different ages. In multiple linear regressions adjusted for race, marital status, parity, adequacy of prenatal care, and urban/rural residence, the decrement in mean birth weight associated with smoking grew steadily from 117 g for the youngest mothers (age less than 16 years) to 376 g for the oldest (age 40 years or more). Similarly, the adjusted relative risk of having a low birth weight (less than 2,500 g) for smokers compared with nonsmokers was lowest for mothers aged 16–17 years, at 1.43 (95% confidence interval 1.22–1.68), and increased steadily to 2.63 (95% confidence interval 1.77–3.90) for mothers aged 40 or more. This result suggests that the effect of exposure to cigarette smoking during pregnancy is modified by advancing maternal age. Further research using data that more precisely measure the exposure (cigarettes per day, years smoked) could help further clarify this issue and better address the public health question of whether smoking cessation programs ought to focus limited resources more selectively toward pregnant smokers in particular age groups. Am J Epidemiol 1994;139:1008–15.
Low-Dose Oral Contraceptive Use and the Risk of Myocardial Infarction

Lynn Rosenberg, ScD; Julie R. Palmer, ScD; R. Sowmya Rao, MS; Samuel Shapiro, MB, FRCP(Edin)

**Background:** Studies of oral contraceptives (OCs) containing 50 μg or more of estrogen suggest an increased risk of myocardial infarction (MI) among current users, particularly if they smoke heavily.

**Objective:** To assess whether use of the newer lower-dose OCs increases the risk of MI.

**Methods:** A case-control study was conducted from January 1985 through March 1999 in 75 hospitals in the greater-Boston and greater-Philadelphia areas. Data on OC use and MI risk factors were obtained by interview from 627 women with a nonfatal first MI (cases) and 2947 female hospital controls younger than 45 years.

**Results:** The overall odds ratio (OR) for current OC use relative to never used was 1.3 (95% confidence interval [CI], 0.8-2.2). The OR was elevated, 2.5 (95% CI, 0.9-7.5), among heavy smokers (≥25 cigarettes per day) but close to 1.0 among lighter smokers (OR = 0.8) and nonsmokers (OR = 1.3). For current OC use together with heavy smoking relative to nonuse and nonsmoking, the OR was 32 (95% CI, 12-81), considerably greater than that for heavy smoking alone, 12 (95% CI, 8.6-16). The ORs did not vary according to the type of formulation or the dose of estrogen; there were too few users to assess the new 20-μg preparations. Past OC use was unrelated to risk.

**Conclusion:** Current use of low-dose OCs in the United States is unrelated to an increased risk of MI among non-smokers and light smokers, but users who smoke heavily may be at greatly increased risk.

_Arch Intern Med. 2001;161:1065-1070_
If HRT is used soon after menopause, it appears protective for CHD.

If HRT is used years after menopause, It appears to be a risk factor for CHD.
Certain anti-hypertensives do not work well in black patients (race is an effect modifier)
Certain anti-hypertensives appear to work better in black patients (race is an effect modifier).
"Cause I gots dat hypertizzle all up in my cardizzle."

"I ain't gonna try it... You try it!! No way!! I ain't gonna try it!!"

"Well, somebody's gotta try it!!"

"Well, that somebody ain't gonna be me!!"

"BIDIL, THE FIRST GOVERNMENT APPROVED DRUG SPECIFICALLY MADE FOR A PARTICULAR RACE, IS SUPPOSED TO HELP BLACK FOLKS COMBAT HEART FAILURE..."

"NOW... GIVEN THE UNITED STATES GOVERNMENT'S LESS THAN STELLAR TRACK RECORD CONCERNING THE HEALTH OF ITS BLACK CITIZENS...

"... IT'S REALLY NO SURPRISE THAT FOLKS MAY BE A BIT RELUCTANT TO TRY THE DRUG..."

"DO YOU REMEMBER THE LAST DRUG THE U.S. GOVT. TARGETED TOWARDS BLACKS?"

"TOO BAD THIS STUFF DIDN'T COME OUT BACK WHEN I WAS IN COLLEGE...."

"WHERE, Y'KNOW. TRY THESE... SURE!! WHAT ARE THEY??"

"CUZ I TOOK ANYTHING PEOPLE HANDED ME..."

"I CAN'T WAIT TO HEAR THE APOLOGY THE SENATE WILL GIVE 50 YEARS FROM NOW FOR THE INFAMOUS BIDIL 'MISTAKE'."

"THE U.S. SENATE HEREBY APOLOGIZES TO THE LAST SURVIVING BLACK PERSON..."
Ethnic-Specific Differences in Bronchodilator Responsiveness Among African Americans, Puerto Ricans, and Mexicans with Asthma

Mariam Naqvi, B.S., Shannon Thyne, M.D., Shweta Choudhry, Ph.D., M.Sc., Hui-Ju Tsai, Ph.D., Daniel Navarro, M.D., Richard A. Castro, M.D., Sylvette Nazario, M.D., Jose R. Rodriguez-Santana, M.D., Jesus Casal, M.D., Alfonso Torres, M.D., Rocío Chapela, M.D., H. Geoffrey Watson, M.D., Kelley Meade, M.D., Michael LeNoir, M.D., Pedro C. Avila, M.D., William Rodriguez-Cintron, M.D., and Esteban González, Burchard, M.D., M.P.H.

1University of California, San Francisco, California, USA
2San Juan VAMC, University of Puerto Rico School of Medicine, San Juan, Puerto Rico
3Pediatric Pulmonary Program of San Juan, San Juan, Puerto Rico
4Instituto Nacional de Enfermedades Respiratorias (INER), Mexico City, Mexico
5The James A. Watson Wellness Center, Oakland, California
6Children’s Hospital and Research Institute, Oakland, California
7Bay Area Pediatrics, Oakland, California
8Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

Socioeconomic and environmental differences do not fully explain differences in asthma prevalence, morbidity, and mortality among Puerto Ricans, African Americans, and Mexican Americans. Differences in response to albuterol may be a factor. We compared bronchodilator responsiveness between these three populations. All groups demonstrated below expected responsiveness. Puerto Ricans of all ages and African American children with moderate-to-severe asthma demonstrated the lowest responsiveness overall. Among subjects with moderate-to-severe asthma, children were even less likely than adults to show the expected bronchodilator response. We conclude that ethnic-specific differences in bronchodilator drug responsiveness exist between Mexicans, Puerto Ricans, and African Americans with asthma. This may be of importance in asthma management.
I Am a Racially Profiling Doctor

By Sally Satel
Published: May 5, 2002

In practicing medicine, I am not colorblind. I always take note of my patient’s race. So do many of my colleagues. We do it because certain diseases and treatment responses cluster by ethnicity. Recognizing these patterns can help us diagnose disease more efficiently and prescribe medications more effectively. When it comes to practicing medicine, stereotyping often works.
Is “personalized medicine” the ultimate example of effect modification?

**Figure 1** Pharmacogenomic approach to personalized medicine. Drug therapy is chosen for each patient based on their particular genetic profile.
23andMe provides you with genetic information, but does not sequence your entire genome or perform predictive or diagnostic tests. Rather, we use currently available technology to examine your DNA sequence at a large number of variable sites called SNPs. Since this SNP information is difficult to interpret on its own, we review the most up-to-date biomedical literature on genetic associations and provide you your genotype information in the context of current scientific knowledge.

While we may be able to tell you that researchers have found your particular genotype to be associated with an increased chance of developing a particular condition, we cannot tell you whether you actually have a specific disease, or whether you will develop a specific disease in the future.

There are several reasons why we cannot provide diagnoses or otherwise assess your health. First, because we don't sequence your entire genome, we may miss variation that has an impact on your health. Genetic testing services, which restrict themselves to a relatively small set of diseases, provide more exhaustive analysis of the relevant genes. More importantly, in order to make a diagnosis, your doctor considers not only your genetic information, but also your particular personal and family history and your physical condition, as well as any symptoms you are experiencing. Other confirmatory tests are usually required, since your genotype is only part of the equation. If you learn that your personal genetic information suggests that you have a higher than average chance of developing a particular disease, you may wish to discuss your genetic information with your physician or another medical expert.
Male circumcision has been shown to protect men from acquiring HIV infection during sex with women — it has reduced female-to-male transmission rates by 48% to 60% in sub-Saharan Africa — but that protective effect appears less reliable among men who have sex with men, according to a new meta-analysis published Oct. 7 in the *Journal of the American Medical Association (J.A.M.A.)*.

**Circumcision appears to have a protective effect in heterosexual men, but not homosexual men**

---

**New Debate Over Circumcision, HIV Reduction**

*By CATHERINE GUTHRIE  Tuesday, Oct. 07, 2008*

Lenora Gim / Photonica / Getty

**Heterosexual**

Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial

**Men who have sex with men**

Circumcision Status and Risk of HIV and Sexually Transmitted Infections Among Men Who Have Sex With Men

A Meta-analysis

---

**Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial**

**Methods and Findings**

Randomized, controlled intervention trials of male circumcision have shown that male circumcision reduces the risk of acquiring HIV infection. This trial randomized 2,826 men into two arms: one received male circumcision, and the other received a sham procedure. The results showed that male circumcision reduced the risk of acquiring HIV infection by 60% (95% CI: 0.36-0.78). This finding is supported by observational studies of circumcision, which have consistently shown a lower risk of acquiring HIV infection among circumcised men compared to uncircumcised men.
Heterogeneity of effects

Can occur at the level of:

- Individual study: within subgroups of a single study or trial
  - Seen in subgroup or stratified analyses within a study

- Across studies: if several studies are done on the same topic, the effect measures may vary across studies
  - Seen in meta-analyses (across trials)
Heterogeneity of effects within a trial or study

- The GISSI trial showed that streptokinase reduced overall mortality roughly 20%. Subgroup analyses suggested that benefit was confined to patients with anterior myocardial infarction, to those under the age of 65 years, and to those treated within 6 hours of the onset of symptoms. But power in each subgroup was low.
  - Subsequent studies demonstrated benefit irrespective of site of infarction, age of patient, and time from onset of symptoms to treatment.

- ISIS-2 trial on streptokinase and aspirin: investigators presented results by the astrological sign under which patients were born. Aspirin was clearly beneficial overall and for persons born under all signs except Libra and Gemini, for which apparent harmful effects were observed.
Hazards of subgroup analyses

- When multiple interaction tests are conducted, each using a nominal criterion (say, P=0.05) to assess statistical significance, the probability of a false positive result — that is, of appearing to find an interaction when none exists — can be greatly inflated.
- For example, when treatments have identical efficacy, the probability of finding at least one "statistically significant" interaction test when 10 independent interaction tests are undertaken is 40 percent.

![Graph showing the probability of at least one false positive result with varying numbers of subgroup tests.]

### 3: Probability of at least one significant result at the 5% significance level given no true differences

<table>
<thead>
<tr>
<th>Number of tests</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
</tr>
<tr>
<td>3</td>
<td>0.14</td>
</tr>
<tr>
<td>5</td>
<td>0.23</td>
</tr>
<tr>
<td>10</td>
<td>0.40</td>
</tr>
<tr>
<td>20</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Cook, MJA, 2004

Lagakos, NEJM 2006
A consumer's guide to subgroup analyses

Guidelines for deciding whether apparent differences in subgroup response are real

1. Is the magnitude of the difference clinically important?
2. Was the difference statistically significant?
3. Did the hypothesis precede rather than follow the analysis?
4. Was the subgroup analysis one of a small number of hypotheses tested?
5. Was the difference suggested by comparisons within rather than between studies?
6. Was the difference consistent across studies?
7. Is there indirect evidence that supports the hypothesised difference?

Heterogeneity in effects across studies (meta-analyses)

Association between smoking and TB mortality

Figure 5. Forest plot of studies\textsuperscript{29-33} that examined smoking and tuberculosis mortality. The sex and age of the study population are shown on the y-axis.
Heterogeneity in effects across studies (meta-analyses)

Meta-analysis on efficacy of BCG vaccination for TB

<table>
<thead>
<tr>
<th>Trial (Latitude)</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Madanapaleo</td>
<td>0.80 (0.52,1.25)</td>
</tr>
<tr>
<td>Madras</td>
<td>1.01 (0.89,1.14)</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>0.71 (0.57,0.89)</td>
</tr>
<tr>
<td>Haiti</td>
<td>0.20 (0.08,0.50)</td>
</tr>
<tr>
<td>South Africa</td>
<td>0.63 (0.39,1.00)</td>
</tr>
<tr>
<td>Georgia</td>
<td>1.56 (0.37,6.53)</td>
</tr>
<tr>
<td>Georgia</td>
<td>0.98 (0.58,1.66)</td>
</tr>
<tr>
<td>Chicago</td>
<td>0.26 (0.07,0.92)</td>
</tr>
<tr>
<td>Chicago</td>
<td>0.25 (0.15,0.43)</td>
</tr>
<tr>
<td>Northern USA</td>
<td>0.46 (0.39,0.54)</td>
</tr>
<tr>
<td>Northern USA</td>
<td>0.41 (0.13,1.26)</td>
</tr>
<tr>
<td>UK</td>
<td>0.24 (0.18,0.31)</td>
</tr>
<tr>
<td>Canada</td>
<td>0.20 (0.09,0.49)</td>
</tr>
</tbody>
</table>

Fig 4. Forest plot of trials of BCG vaccine to prevent tuberculosis. Trials are ordered according to the latitude of the study location, expressed as degrees from the equator. No meta-analysis is shown (CI = confidence intervals, RR = relative risk) (adapted from Colditz et al.47).
Subgroup analysis within meta-analysis

Beta-carotene intake and cardiovascular mortality

Definition based on the comparison between observed and expected joint effects of a risk factor and a third variable [deviation from additive or multiplicative joint effects]

Is the whole more (or less) than the sum (or product) of its parts?

This is often called “statistical interaction”
Observed vs expected joint effects of a risk factor and a third variable

A. When there is no interaction, the observed joint effect of risk factors A and Z equals the sum of their independent effects:

\[ A + Z = A + Z \]

No interaction

B. When there is positive interaction (synergism), the observed joint effect of risk factors A and Z is greater than the expected on the basis of summing their independent effects:

\[ A + Z = A + Z * \]

Positive interaction

* Excess due to positive interaction

C. When there is negative interaction (antagonism), the observed joint effect of risk factors A and Z is smaller than the expected on the basis of summing their independent effects:

\[ A + Z = A + Z \uparrow \]

Negative interaction

† “Deficit” due to negative interaction
Definition based on the comparison between observed and expected joint effects of a risk factor and a third variable

- Interaction on an “additive” scale (additive interaction)
  - Effect measure modification when risk difference is used as measure of effect
  - Additive statistical model:
    - Linear regression: \( y = a + b_1x_1 + b_2x_2 \)

- Interaction on a “multiplicative” scale (multiplicative interaction)
  - Effect measure modification when risk ratio is used as measure of effect
  - Multiplicative statistical model:
    - Logistic regression: \( \frac{p}{1-p} = e^{b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_kx_k} \)
Example: Smoking, asbestos, lung cancer
Example: Smoking, asbestos & lung cancer

Death rates from lung cancer (per 100,000)

<table>
<thead>
<tr>
<th>Cigarette smoking</th>
<th>Asbestos exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
<td>11</td>
</tr>
<tr>
<td>Yes</td>
<td>123</td>
</tr>
</tbody>
</table>

Does smoking modify the effect of asbestos on cancer?

Risk difference in non-smokers = 47 (58 – 11)
Risk difference in smokers = 479 (602 – 123)
Risk ratio in non-smokers = 5.2 (58/11)
Risk ratio in smokers = 4.9 (602/123)

Data: Hammond, 1979
Consider a study to explore the association between age and incidence of a disease.

Question: is the association between age and disease modified by sex?
When data are stratified (by sex):

Question: is the association between age and disease modified by sex?

Rothman, 2002
Answer: depends on the scale used!

E.g.

90% - 85% = 5%
90/85 = 1.05

Rate difference stays constant; Rate ratio decreases

Incidence

E.g.

10% - 5% = 5%
10/5 = 2

Age

Rothman, 2002
What if the lines were like this:

E.g.

100% - 50% = 50%
100/50 = 2

Rate ratio stays constant; Rate difference increases

E.g.

10% - 5% = 5%
10/5 = 2

Incidence rate

Age

Different Slopes for Different Folks!

Rothman, 2002
Statistical interaction is scale-dependent!

- When interaction is absent using ratio measures, it will necessarily be present when risk difference measures are used, and *vice versa*
- Because interaction is “scale-dependent” the term “effect measure modification” is more specific than “effect modification”
  - It’s important to specify which scale (risk difference vs. risk ratio) was used in the analysis
Additive Interaction: departure from an additive statistical model

Death rates from Lung cancer (per 100,000)

<table>
<thead>
<tr>
<th>Cigarette smoking</th>
<th>Asbestos exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>No</td>
<td>11 (baseline risk)</td>
</tr>
<tr>
<td>Yes</td>
<td>123</td>
</tr>
</tbody>
</table>

- Excess risk due to smoking: $123 - 11 = 112$
- Excess risk due to asbestos: $58 - 11 = 47$
- Excess risk expected due to both (under + model): $112 + 47 = 159$
- Total observed excess risk: $602 - 11 = 591$ !!

Observed excess risk is much higher than what we expect from our additive model: there is interaction (on additive scale)!

Data: Hammond, 1979
Example: Smoking & Asbestos

Death rates from Lung cancer (per 100,000)

Data: Hammond, 1979
*Note that when the independent relative odds for A and Z are added, the baseline is added twice; thus, it is necessary to subtract 1.0 from the expected joint OR: that is, Expected OR_{A+Z+} = (Excess due to A + baseline) + (Excess due to Z + baseline) - baseline = OR_{A+Z-} + OR_{A-Z+} - 1.0.

**Figure 6-3** Schematic representation of the meaning of the formula, Expected OR_{A+Z+} = Observed OR_{A+Z-} + Observed OR_{A-Z+} - 1.0.
Multiplicative Interaction: departure from a multiplicative statistical model

Death rates from Lung cancer (per 100,000)

<table>
<thead>
<tr>
<th>Cigarette smoking</th>
<th>Asbestos exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>123</td>
</tr>
</tbody>
</table>

- RR due to smoking: \(\frac{123}{11} = 11.2\)
- RR due to asbestos: \(\frac{58}{11} = 5.3\)
- RR expected due to both (under x model): \(11.2 \times 5.3 = 59.4\)
- Total observed RR: \(\frac{602}{11} = 54.7\)

Observed RR is close to what we expect from our multiplicative model: this is **NO** interaction on a multiplicative scale

Data: Hammond, 1979
Male TB patients who were both smokers and alcoholics had a higher RR than those who were either only smokers or only alcoholics.

Is there interaction on a multiplicative scale?
Additive or multiplicative model?

The additive model underpins the methods for assessing biological interaction (causal pie model by Rothman)
- Interaction here means a departure from additivity of disease rates (risk difference is the key measure)
- Some believe that risk difference scale is of greatest public health importance (because it’s based on AR and PAR)

In contrast, many of the models used in epi analyses are inherently multiplicative (e.g. logistic regression)
- Vast majority of epi analyses are based on a multiplicative model and hence most epi studies implicitly use the multiplicative scale (risk ratio is the key measure)
- This is because most epi studies report RR and OR estimates and use regression models such as logistic and survival analyses – these models inherently use ratio measures and are therefore multiplicative

Ahlbom A et al. Eur J Epi 2005
Regardless of the scale, why is interaction/effect modification important?

- Better understanding of causation
  - e.g. smoking and asbestos; diet and PKU

- Identification of “high-risk” groups
  - e.g. influenza can lead to serious complications in specific groups: young, elderly, and those with chronic diseases
  - e.g. women who smoke heavily and use OC are at high risk for myocardial infarction
  - e.g. TB patients who smoke and drink are at high risk for mortality

- Target interventions at specific subgroups
  - e.g. flu vaccines are usually given to only specific groups – aged 65 or older
  - e.g. best time to give measles vaccine is 12 – 15 months
  - e.g. circumcision for heterosexual men
Readings

- Rothman text:
  - Chapter 9: Measuring Interactions

- Gordis text:
  - Chapter 15
What to call your Professor

Have they said you can call them by their first name?

No

Do they cringe when you call them by their first name?

Yeah

My Prof. is totally cool.

Uh, a little

Then they didn’t really mean it. Start over.

Less than 4 years

More than 4 years

(it’s kind of irrelevant)

Have you gotten drunk with him/her at a conference?

Yes

No

Are they from California?

Yes

Dude...

No

Can you think of a word that’s less formal than “Prof./Dr.” but not as disrespectful as using their first name (kind of like “Dad” with your father)?

No

Neither can we. Proceed with extreme caution.

Hey there, um, ...?

Congrats! You’re on a first-name basis with your Professor!

If you haven’t noticed, it’s probably you (grr)

#5@% Golden Boy/Girl!!

Wait, are you an undergrad? Undergrads must never call Professors by their first name. It’s just weird.