Week 2 – Thursday Session

Effect Measure Modification

EPI202 – Epidemiologic Methods II
Murray A. Mittleman, MD, DrPH
Department of Epidemiology, Harvard TH Chan School of Public Health

HARVARD T.H. CHAN
SCHOOL OF PUBLIC HEALTH

Week 2: Discussion Topics

1. Stratified person-time data
   - Bias vs. efficiency
   - Notation
   - Hypothesis tests
   - Point and interval estimates – weighted averages
   - Assumption of homogeneity
2. Stratified count data
   - Hypothesis tests
   - Point and interval estimates – weighted averages
   - Assumption of homogeneity
3. Effect Measure modification
   - Definition
   - Impact on generalizability
   - Scale dependence
   - The H statistic for difference and ratio measures
   - Relative excess risk due to interaction

Effect Measure Modification

- In the presence of effect measure modification, the magnitude of the association between exposure and disease varies according to the value of (across strata of) a third factor, which is called an effect modifier.
- Effect measure modification is an intrinsic phenomenon and cannot be eliminated from a study through clever design
- Effect measure modification is a finding to be reported rather than a bias to be avoided
- Synonyms: interaction, synergy, antagonism

External Validity

- The possibility of **effect measure modification** reduces the **external validity** (generalizability or transportability) of a study which is restricted to a particular sub-population

- When the results of a study are generalizable to other, larger populations, we say that this study is **externally valid**.

In the data shown, there appears to be effect measure modification on

	Age < 65		Age 65+	
	M	F	M	F
Deaths	14	10	76	121
Person-years	1,516	1,701	949	2245
$\hat{I}/10^3$	9.2	5.9	80.0	53.9
ÎRR	1.6		1.5	
$\hat{I}RD/10^3$ PY	3.3		26.1	

- The multiplicative scale
- The additive scale ✓
- Both the multiplicative and additive scales
- There does not appear to be effect measure modification

Total Results: 0

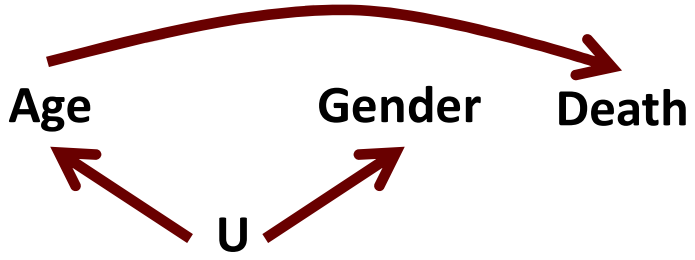
Based on the data below including the relationship between age and sex in these data, age appears to be

	Age < 65		Age 65+	
	M	F	M	F
Deaths	14	10	76	121
Person-years	1,516	1,701	949	2245
$\hat{I}/10^3$	9.2	5.9	80.0	53.9
ÎRR	1.6		1.5	
$\hat{I}RD/10^3$ PY	3.3		26.1	

- A confounder
- A modifier on the multiplicative scale
- A modifier on the additive scale
- I'm not sure

Total Results: 0

Confounding vs. Effect Modification (IRR)
Age, Gender and Mortality Following Trigeminal Neuralgia



	Age < 65	Age 65+	Crude
$\hat{IRR}$	1.6	1.5	1.1

- Age is a confounder of the gender-related differential mortality among this population.

Test: 1. Age and Gender in the study base; 2. Age and Death among the non-exposed
If associations exist in these two links, age should be a confounder between gender and death.

- Age is not an effect modifier on the multiplicative scale.

Test of Homogeneity (IRR)
Null and Alternative Hypotheses

- H_0: The rate ratio is the same across all I levels of the stratification variable(s)
  - ↔ There is no effect modification of the IRR by the stratification variable(s)
  - ↔ The rate ratio is *homogeneous* across the strata
  - ↔ $IRR_1 = IRR_2 = \ldots = IRR_I$
  - ↔ $IRR_i = IRR_j$ for all i,j

- H_A: The rate ratio is not the same across all I levels of the stratification variable(s)
  - ↔ There is effect modification of the IRR by one (or more) of the stratification variable(s)
  - ↔ The rate ratio is *heterogeneous* across the strata
  - ↔ At least one of the IRRs does not equal at least one of the others
  - ↔ $IRR_i \neq IRR_j$ for at least one i,j pair

Test of Homogeneity (IRR)
Form of the Test Statistic

- In general, tests for homogeneity of ratio measures, i.e. for effect modification on the multiplicative scale, have the following form

$$H = \sum_{i=1}^{I} \frac{[ln(\hat{IRR}_i) - ln\,\hat{IRR}_{MH}]^2}{\hat{Var}_i[ln(\hat{IRR}_i)]} \sim \chi^2_{I-1}$$

Degree of freedom:
strata - 1

- Where :
  - $\hat{IRR}_i$ is the stratum specific estimate of the incidence rate ratio
  - $\hat{IRR}_{MH}$ is the MH summary estimate of the incidence rate ratio, assuming homogeneity across all strata, and
  - $\hat{Var}[ln(\hat{IRR}_i)] = \frac{1}{a_i} + \frac{1}{b_i}$ is the (large-sample) variance of the stratum specific estimate of the *ln* incidence rate ratio.

Test of Homogeneity (IRR)
Test Statistic Computation

	Age < 65	Age 65+
$\hat{IRR}_i$	1.57	1.49
$ln(\hat{IRR}_i)$	0.452	0.396
$\hat{Var}[ln(\hat{IRR}_i)]$	0.171	0.021
IRR_{MH}	1.50	
$ln(IRR_{MH})$	0.405	

$$H = \sum_{i=1}^{I} \frac{[ln(\hat{IRR}_i) - ln\,\hat{IRR}_{MH}]^2}{\hat{Var}_i[ln(\hat{IRR}_i)]}$$

$$= \frac{(0.452-0.405)^2}{0.171} + \frac{(0.396-0.405)^2}{0.0.021} = 0.016$$

Test of Homogeneity (IRD)
Null and Alternative Hypotheses

- H_0: The rate difference is the same across all I levels of the stratification variable(s)
  - ↔ There is no effect modification of the IRD by the stratification variable(s)
  - ↔ The rate difference is *homogeneous* across the strata
  - ↔ $IRD_1 = IRD_2 = \ldots = IRD_I$
  - ↔ $IRD_i = IRD_j$ for all i,j

- H_A: The rate difference is not the same across all I levels of the stratification variable(s)
  - ↔ There is effect modification of the IRD by one (or more) of the stratification variable(s)
  - ↔ The rate difference is *heterogeneous* across the strata
  - ↔ At least one of the IRDs does not equal at least one of the others
  - ↔ $IRD_i \neq IRD_j$ for at least one i,j pair

Test of Heterogeneity (IRD)
Form of the Test Statistic

- In general, tests for heterogeneity of difference measures, i.e. for effect modification on the additive scale, have the following form

$$H = \sum \frac{\left[I\hat{R}D_i - I\hat{R}D_{summary}\right]^2}{var(I\hat{R}D_i)}$$

- Where :
  - $I\hat{R}D_i$ is the stratum specific estimate of the incidence rate difference
  - $I\hat{R}D_{summary}$ is the summary estimate of the incidence rate difference, assuming homogeneity across all strata, and
  - $var\left(I\hat{R}D_i\right) = \frac{a_i}{N_{1i}^2} + \frac{b_i}{N_{0i}^2}$ is the variance of the stratum specific estimate of the incidence rate difference.

Test of Heterogeneity (IRD)
Test Statistic Computation

	Age < 65	Age 65+
$\hat{IRD}_i$	$3.36/10^3$PY	$26.2/10^3$PY
$\hat{Var}[IRD_i]$	$9.5/10^6$ PY^2	$108/10^6$ PY^2
$IRD_{summary}$	$13.7/10^3$ PY	

$$H = \sum \frac{\left[\hat{IRD}_i - \hat{IRD}_{summary}\right]^2}{var(\hat{IRD}_i)}$$

$$= \frac{\left(3.36/10^3 PY - 13.7/10^3 PY\right)^2}{9.55/10^6 \text{ PY}^2} + \frac{\left(26.2/10^3 PY - 13.7/10^3 PY\right)^2}{108/10^6 \text{ PY}^2} = 12.69$$

Test of Heterogeneity (IRD)
Test Statistic Interpretation

- Degrees of freedom for H = 2-1 = 1
- $Pr[\chi^2_1 > 12.69] = 0.0004$
- In a hypothesis testing framework with a pre-specified 2-sided alpha of 0.05, we reject the null hypothesis of no effect measure modification on the additive scale in these data. This interpretation hinges on the validity of our usual assumption regarding freedom from bias.
- These data are not very consistent with the state of nature described by the null. If the data arose from a single common rate difference, we would only expect to observe this degree of heterogeneity or more once in 25,000 such studies. For this interpretation to be valid, we need to assume no sources of bias.
- This is consistent with our less formal appraisal of the data. The rate difference is nearly tenfold greater in the older patients than in the younger patients, and there is relatively little overlap in the 95% confidence intervals.

Limitations of Tests of Homogeneity

- Tests of homogeneity share the same limitations as other hypothesis tests discussed earlier:
  - they fail to summarize the data with respect to their consistency with any alternative hypotheses
  - nor do they give us any indication of the *power* of the data to detect any alternative hypotheses of interest
- Considerably more data than are often available are needed to detect and characterize effect measure modification, if present.
- Many studies, unless explicitly designed to detect and characterize effect measure modification, will not have sufficient data to do so. Keep in mind that relatively precise stratum-specific estimates are usually required to statistically detect effect measure modification.
- Thus, not rejecting the null hypothesis of no effect measure modification may often be explained by a low statistical power.

Reporting Results

- When effect measure modification is present, it is not generally useful to summarize over strata with heterogeneous associations.

- Rather, it is of greater interest to report the observed associations by level of the modifier.

- If a summary measure is needed, weights that do not reflect arbitrary features of the study design should be chosen. Instead, ***standardization techniques*** (including ***inverse probability weighting***) should be used, with appropriate population-based weights (more on this later).

Standardization techniques (IPW) is used to compute the summary measure (if EMM exists)
It will give each strata with a weight to adjust.

Can EMM on the Difference Scale be Detected when Estimation is Based on a Multiplicative Model?

- Epidemiologists often estimate associations on the multiplicative scale
  - Multiplicative models are often statistically efficient and convenient
    - Logistic regression
    - Cox Proportional Hazards regression
    - Poisson regression

- Detecting the presence of EMM on the multiplicative scale is straightforward
  - Stratified results
  - Incorporating multiplicative interaction terms in statistical models

- Whether EMM is also present on the additive scale may be of interest for several reasons
  - As an indication of mutual mechanistic interaction under the Rothman sufficient and component cause model
  - To develop intuition about absolute impact of exposure in subsets of the population

Can EMM on the Difference Scale be Detected when Estimation is Based on a Multiplicative Model?

- The Relative Excess Risk due to Interaction (RERI) is one metric that allows the assessment of EMM on the additive scale from multiplicative parameters

Relative Excess Risk due to Interaction (RERI)

Consider two exposures, A1 and A2:

	$A_1=0$	$A_1=1$
$A_2=0$	CI_{00}	CI_{10}
$A_2=1$	CI_{01}	CI_{11}

If there is no EMM on the additive scale then
$CI_{10} - CI_{00} = CI_{11} - CI_{01}$ and therefore, $CI_{11} - CI_{01} - CI_{10} + CI_{00} = 0$

Divide each term by CI_{00} and no EMM on the additive scale implies
$CIR_{11} - CIR_{01} - CIR_{10} + 1 = 0$.

This expression is called the Relative Excess Risk due to Interaction (RERI). The RERI allows the use of multiplicative parameters to determine whether there is EMM on the additive scale.

RERI = 0 if there is no EMM on the additive scale
RERI > 0 if the absolute effect of A1 is greater in the presence of A2 (and vice versa)
RERI < 0 if the absolute effect of A1 is weaker in the presence of A2 (and vice versa)

Relative Excess Risk due to Interaction (RERI)

10-year Cumulative Incidence Ratio of Lung Cancer
by Smoking and Asbestos Status in Telemark, Norway

	No Asbestos	Asbestos
Non-smoker	1	6.09
Smoker	8.64	40.91

$$RERI = CIR_{11}\text{-}CIR_{10}\text{-}CIR_{01}+1$$
$$RERI = 40.91 - 8.64 - 6.09 + 1$$
$$RERI = 27.18$$

- Using only results on the multiplicative scale
  - The RERI indicates that there is EMM on the additive scale
  - Asbestos is more harmful among smokers than non-smokers and
  - Smoking is more harmful among those exposed to asbestos than those not

Week 2 Data Exercise

HAVE A GOOD WEEKEND