

Epidemiology in Occupational Medicine

OEMAC Pre-Conference 2015
Organized by the Canadian Board of Occupational Medicine
September 27, 2015

Ron House MD, MSc, FCBOM, FRCPC



St. Michael's
Inspired Care. Inspiring Science.

Disclosures

No conflicts of interest to declare

- Financial support
- In-kind support
- Other potential sources of COI

Learning Objectives

- Describe the main study designs used in occupational epidemiological studies
- Identify the key sources of bias and random error in occupational epidemiology studies
- Apply this information to the critical appraisal of the occupational epidemiological literature

Occupational Epidemiology: Definition

- The application of epidemiologic methodology to the study of the risk of health outcomes due to workplace exposures and interventions
- Related to:
 - public health epidemiology
 - clinical epidemiology

Occupational Epidemiology: Exposures

- Traditional:
 - job title
 - work location
- Newer:
 - occupational hygiene measurements
 - job exposure matrices
 - workplace organization/stress
 - biomarkers / genetic markers

Occupational Epidemiology: Outcomes

- Traditional outcomes:
 - mortality
 - cancer mortality / incidence
- Newer outcomes:
 - morbidity:
 - Multiple chemical sensitivity
 - Chronic CNS effects of solvents
 - Repetitive strain injury
 - Neurological component of HAVS
 - problems:
 - Case definition
 - Measurement

Possible Interpretations of Associations Between Exposure and Outcome in Epidemiologic Studies

- True association
 - Systematic error (bias):
 - Refers to anything in a study that produces results which deviate systematically from the truth
 - Confounding
 - Measurement bias
 - Selection bias
 - Random error (chance):
 - type I (α)
 - type II (β)
- Power = 1 - β

Relation between Exposure and Outcome may be Influenced by Other Variables

- Confounding Factors
- Effect Modifiers

Confounding

- Definition:
 - A determinant of the outcome
 - Associated with the exposure under investigation
 - Not an intermediate step in the causal pathway between exposure and outcome
- Example:
 - Smoking in workers exposed to an occupational lung carcinogen

Stratified Analysis

Stratifying variable (ie. smoking)

		Outcome	
		Yes	No
<i>Smokers</i>			
Exposure	Yes	A ₁	B ₁
	No	C ₁	D ₁
<i>Non-Smokers</i>			
Exposure	Yes	A ₂	B ₂
	No	C ₂	D ₂

Confounding

<u>Group</u>	<u>RR (hypothetical values)</u>
Overall (with confounding)	4.0
Stratified by Confounder:	
Level 1	2.5
Level 2	2.5

The relative risks do not differ in the two strata
(aside from random variability in measuring RR)

Control of Confounding

- Restriction (entry criteria)
- Matching
- Randomization in RCT
- Analysis:
 - Stratified analysis (Mantel-Haenszel)
 - Multivariable analysis (e.g multiple linear regression, logistic regression, ANOVA)

Effect Modification

- Effect Modifier

- The risk of the outcome due to the exposure depends on the level of a third factor (the effect modifier)
- Modifies the effect of a causal factor

- Example: the risk of lung cancer from some carcinogens depends on the type of metabolizing enzyme polymorphism
e.g. CYP1A1 – increased risk of lung cancer from PAH's in cigarette smoke

Effect Modification

<u>Group</u>	<u>RR (hypothetical values)</u>
Overall	4
Stratified by Effect Modifier:	
Level 1	2
Level 2	6

The relative risks differ in the two strata

Evaluation of Effect Modification

- Analysis
 - Stratified Analysis
 - Examine RR in separate strata
 - Multivariable regression
 - Interaction terms (X_1X_2)

Measurement Bias

- Relates to measurement of:
 - exposure
 - outcome
 - confounding factors
 - effect modifiers
- Results in misclassification
 - Non-differential misclassification:
 - exposed/unexposed or those with/without health effect have equal chance of being misclassified
 - produces bias towards null hypothesis (polychotomous variables may sometimes behave differently)
 - Differential misclassification:
 - chance of misclassification not equal among study groups
 - produces bias either toward or away from the null hypothesis (depending on how the misclassification occurs)

Misclassification Example:

Given the following data in a case control study:

		Cases	Controls
Exposure	Yes	100	50
	No	50	100

$$\text{Odds Ratio} = (100 \times 100)/(50 \times 50)$$

$$= 4$$

Misclassification Example (Non-differential)

Now assume 10% of cases and 10% of controls truly not exposed are classified as exposed :

		Cases	Controls
Exposure	Yes	100 + 5 (105)	50 + 10 (60)
	No	50 - 5 (45)	100 - 10 (90)

$$\begin{aligned}\text{Odds Ratio} &= (105 \times 90)/(45 \times 60) \\ &= 3.5 (<4)\end{aligned}$$

Misclassification Example (Differential)

Now, assume 10% of cases truly not exposed are classified as exposed but no misclassification of controls:

		Cases	Controls
Exposure	Yes	100 + 5 (105)	50
	No	50 - 5 (45)	100

$$\begin{aligned}\text{Odds Ratio} &= (105 \times 100)/(45 \times 50) \\ &= 4.7 (>4)\end{aligned}$$

Control of Measurement Bias

- Appropriate measurement methodology
 - Proper instrumentation, calibration, sufficient number of samples
- Standardized assessment
- Blinding

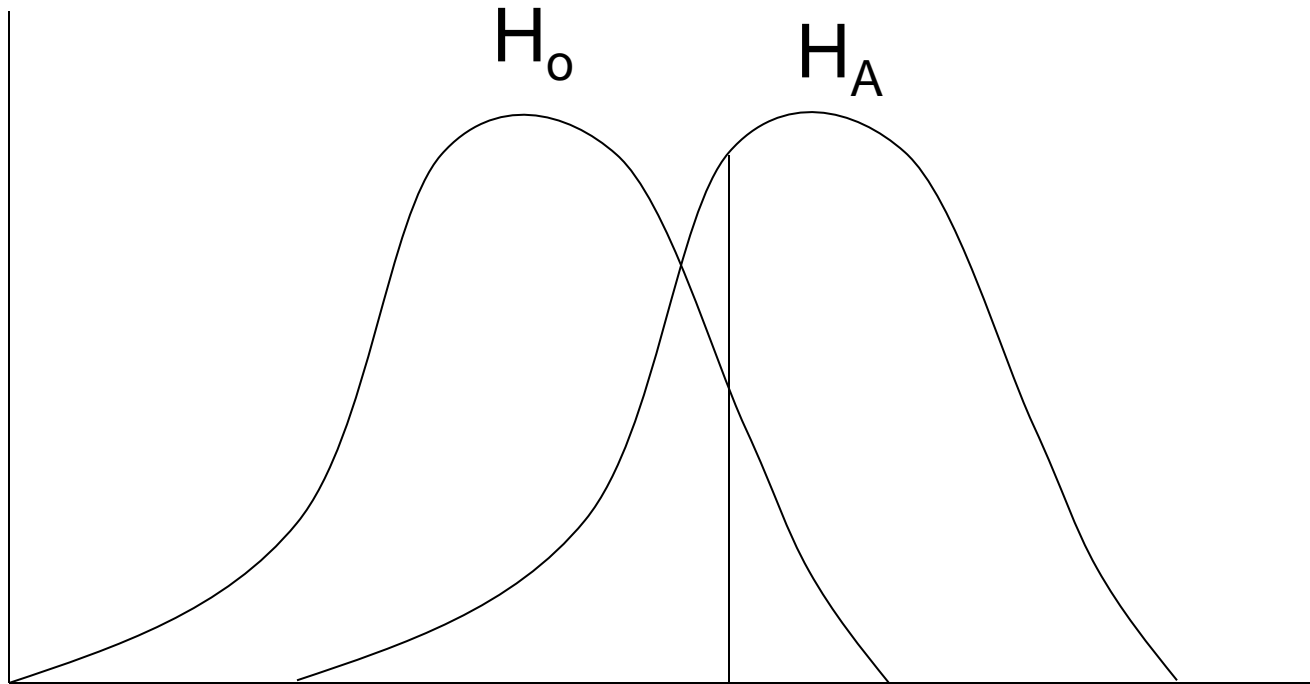
Selection Bias

- **Definition:**
 - the process of selection of the groups to be compared leads to these groups differing with respect to an important determinant of the outcome (other than the exposure of interest)
 - example: Healthy Worker Effect
Volunteers for a wellness program
- **Control:**
 - careful selection of exposed/non-exposed or cases/controls in observational studies
 - randomization for interventional studies

Random Error

- Should consider:
 - Type I Error
 - Type 2 Error

Overlapping Sampling Distributions for the Null and Alternative Hypotheses



Four Possible Outcomes of Hypothesis Testing

Conclusion of test of significance	Truth	
	H_0 True	H_0 False
Do no reject H_0 (not statistically significant)	Correct:(True Negative) H_0 is true, and we do not reject H_0	Type II or beta error: H_0 is false, but we do not reject H_0
Reject H_0 (statistically significant)	Type I or alpha error: H_0 is true, but we reject H_0	Correct: (True Positive) H_0 is false, and we reject H_0

Type I Error

- Should be considered in those studies that report an association to be statistically significant
- Definition: The probability of rejecting the null hypothesis when the null hypothesis is true
- Probability of type I error is increased by:
 - assessment of multiple outcomes
 - multiple subgroup analysis

Type I Error Calculation

- Prob (type I error) = $1 - (1 - \alpha)^N$
 - α = level of significance for each test
 - N = number of separate independent comparisons

Type I Error Example

If SMR's for 20 different causes of death are evaluated (each at an α of 0.05), the overall probability of a type I error in the study is:

$$\begin{aligned} \text{– Probability (type I error)} &= 1 - (1 - \alpha)^N \\ &= 1 - (1 - 0.05)^{20} = 0.64 \end{aligned}$$

Type I Error

- Correction for the type I error:
 - for each comparison the level of significance is reduced:
 - α / n (approximate adjustment)
 - or $1 - (1 - \alpha)^{1/n}$ (exact Bonferroni adjustment)
(for all of the outcomes considered together the probability of a type 1 error is reduced to α)
 $[1 - (1 - \alpha)^{1/n}]^N = \alpha$
 - alternatively, can specify one outcome as hypothesis testing ($\alpha = 0.05$) and the others as hypothesis generating ($\alpha = 0.05$)

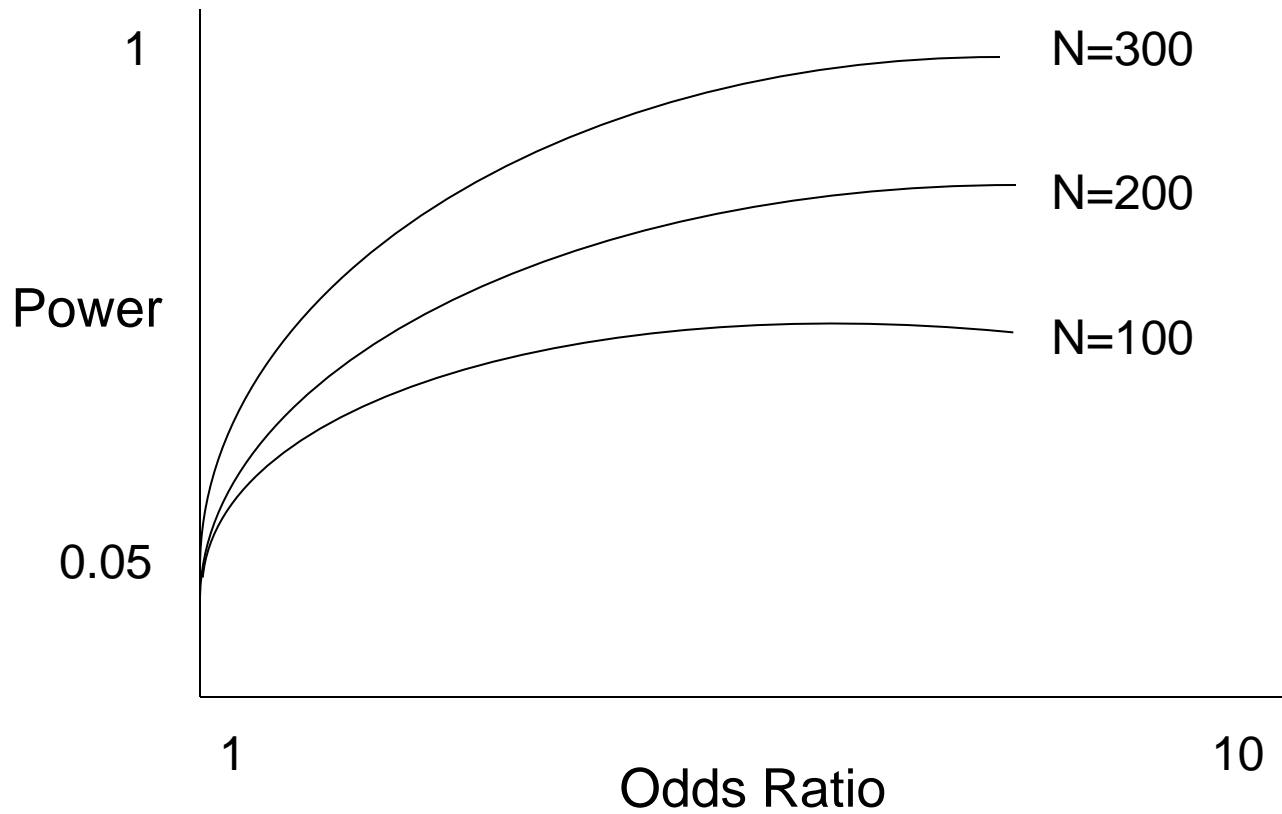
Type 2 Error

- Should be considered in negative studies
- Definition: the probability of failing to reject the null hypothesis when the null hypothesis is false
- Power = 1 – Probability (type 2 error)

Power

- Power of a study depends on:
 - Sample size
 - Effect size (minimum RR you want to be able to detect to be statistically significant)
 - variance of the outcome
 - level of significance
 - tailness (one tail: $Z_{\alpha} = 1.65$; two tail: $Z_{\alpha/2} = 1.96$)

Relation Between Power, Sample Size and Effect Size



Relevance of Power to Critical Appraisal

- Mainly relates to negative studies:
 - If a study is negative it may have had insufficient power to detect the effect due to the exposure to be statistically significant

Validity

Internal validity:

- extent to which the study results are correct for the groups being investigated
- a function of the extent to which **bias** (confounding, selection, measurement) and **random error** have been minimized in a study

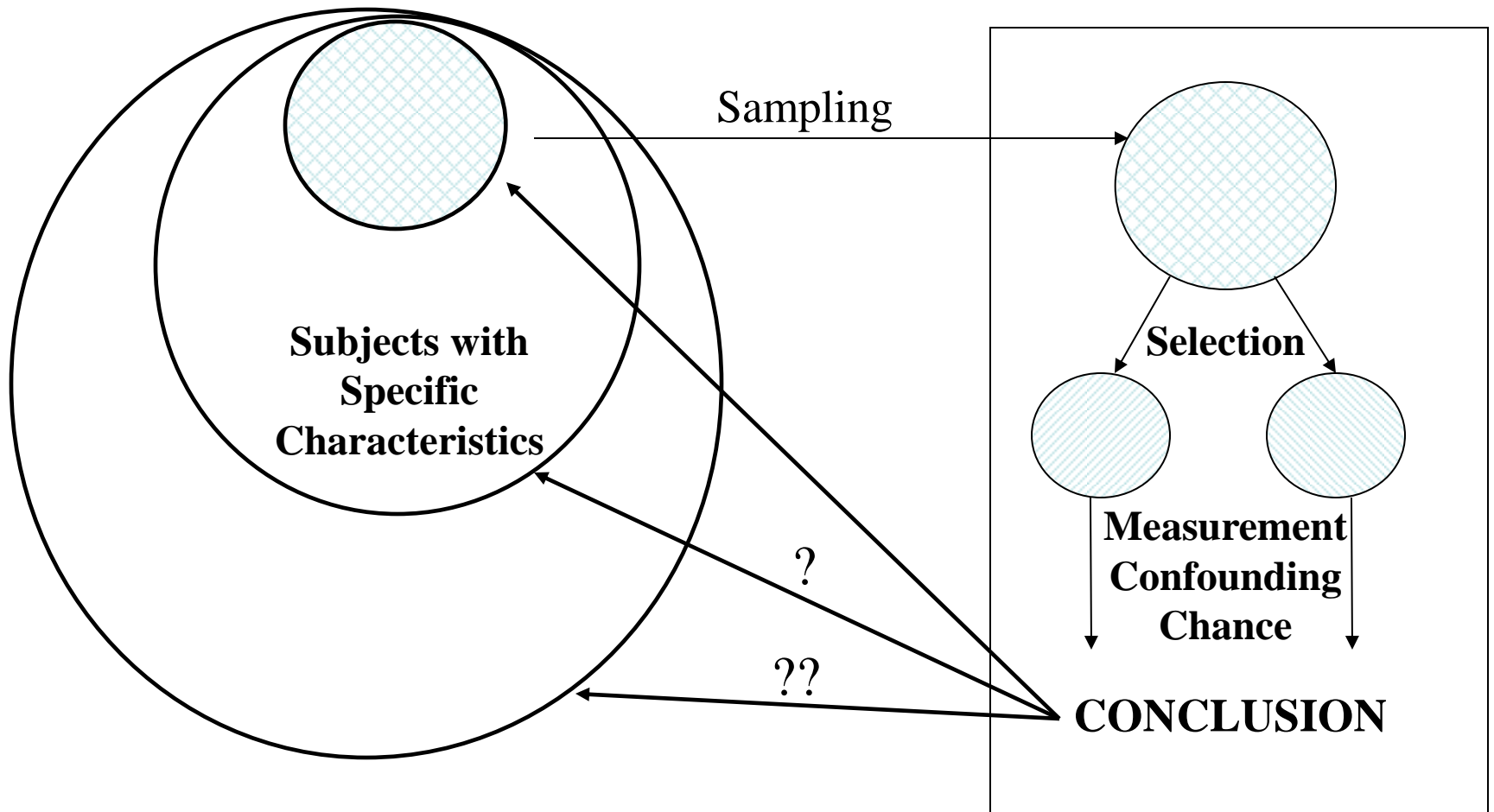
Validity

- External validity:
 - extent to which study results are true in other settings or for other groups; a measure of generalizability
 - affected by recruitment and selection process (study population, inclusion/exclusion criteria and the participation rates)
 - example:
 - lack of occupational studies in women → are results in men generalizable to women?

Relationship Between Internal Validity, External Validity, Bias, and Chance

All Subjects with the Condition of Interest

Internal Validity (The Study)

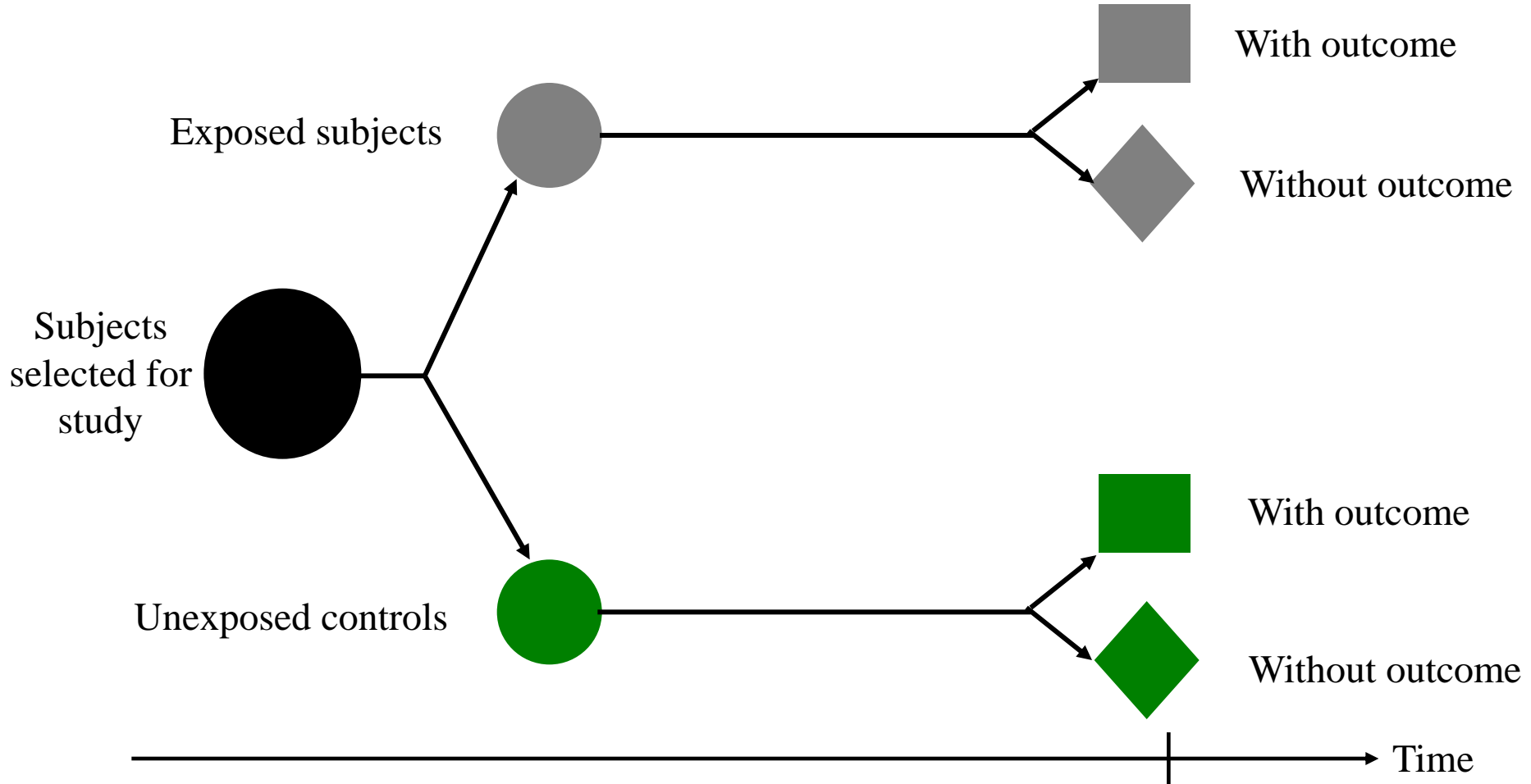


Epidemiologic Study Designs in Occupational Health Research

General Ranking of Quality of Study Designs

1. Experimental studies (RCT)
2. Observational studies:
 - prospective cohort
 - case control
 - historical prospective cohort
 - cross sectional
 - proportional mortality
 - ecologic
3. Descriptive studies

Prospective Cohort Studies



Relative Risk

		Outcome	
		Yes	No
Exposure	Yes	A	B
	No	C	D

- $RR = [A/(A+B)] \div [C/(C+D)]$
- RR tells us how many times more likely exposed individuals are to develop the outcome relative to unexposed individuals
- RR indicates the strength of association
- useful measure for studies of etiology

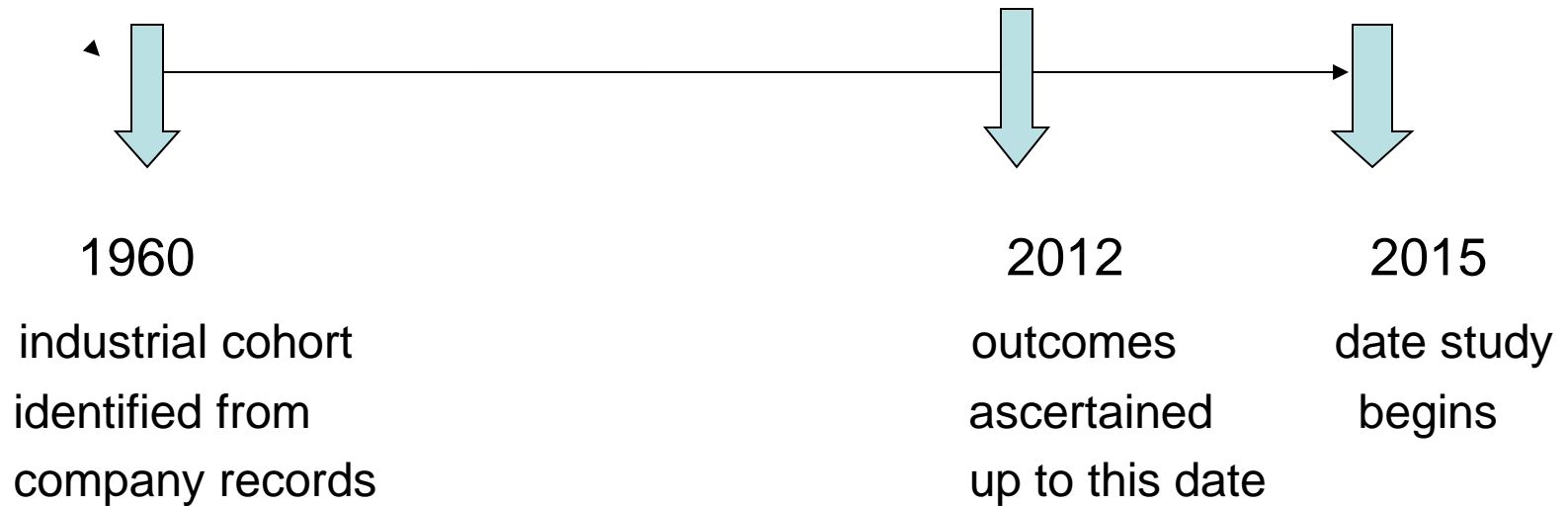
Prospective Cohort Studies

- **Main feature:**
 - exposed and unexposed groups identified and followed prospectively from present
- **Advantages:**
 - can establish temporal relationship between exposure and outcome
 - multiple outcomes can be investigated simultaneously

Prospective Cohort Studies

- **Disadvantages:**
 - if long latency period and/or rare outcome, need to follow very large cohort for long time period (expensive and logistically difficult)
- **Relevance to occupational health:**
 - Used as much as possible (often after several cross sectional studies on the same topic)
 - may be used to investigate outcomes that occur over relatively brief time span.

Retrospective Cohort Design



$$SMR = O/E$$

Person Years Distribution by Age and Calendar Year

2000-04									
1995-99									
1990-94									
1985-89									
1980-84									
1975-79									
1970-74									
1965-69									
1960-64									
	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64

Standardized Mortality Ratio (SMR)

$$\text{SMR} = O/E$$

$$E = \sum (py)_i R_i$$

py_i = number of person years for the i^{th} age, sex, calendar-year stratum in the study group

R_i = outcome rate for the reference group for the i^{th} age, sex, and calendar year stratum

E = number of outcomes expected if the outcome rates in the study group were the same as those in the reference group

Historical Prospective Cohort Studies (Retrospective Cohort Studies)

- **Main feature:**

- choose an exposed group (cohort) that existed in the past and follow it “forward” in historical time to determine outcome(s)
- compare observed versus expected number of outcomes
- expected number of outcome(s) based on rates in reference population and the person years distribution of the cohort

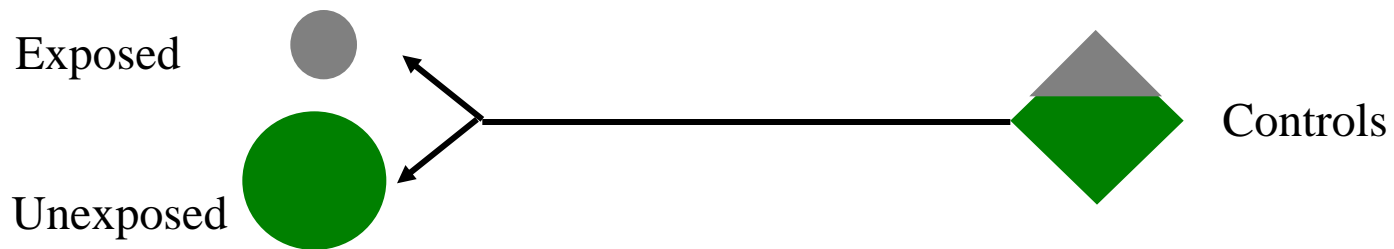
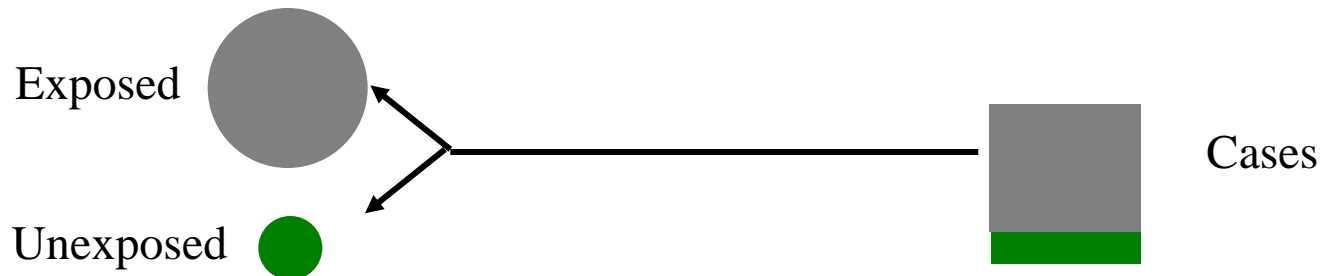
- **Advantages:**

- can follow large cohort for long time period (in historical time)
- good for rare exposures (if large enough industrial cohort can be assembled)
- can examine many outcomes

Historical Prospective Cohort Studies (Retrospective Cohort Studies)

- **Disadvantages:**
 - can only control for a few factors (ie. age, sex, calendar year)
 - restricted to mortality studies and those morbidity studies (e.g. cancer incidence) for which expected numbers can be calculated
 - employment records needed to construct historical cohort may be incomplete or unavailable
- **Relevance to occupational health:**
 - used mainly to investigate mortality, cancer incidence/mortality in industrial cohorts

Case Control Studies



Question: "What happened?"

Onset of Study

Time

A horizontal timeline arrow points to the right, labeled "Time". A vertical tick mark on the timeline is labeled "Onset of Study". The text "Question: 'What happened?'" is positioned to the left of the "Onset of Study" tick mark.

Odds Ratio (OR)

		Outcome	
		Yes	No
Exposure	Yes	A	B
	No	C	D

- Odds Ratio = $[A/C] \div [B/D]$
= AD/BC
- Odds Ratio = ratio of the odds of exposure in the cases to the odds of exposure in the controls

Case Control Studies

- **Main features:**
 - groups are identified on the basis of outcome (cases, controls) and then compared in terms of prior exposure
 - incident cases should be used
 - controls should be obtained from the same source population that produced the cases (i.e. sampling frame for control selection established from the source population that gave rise to the cases)

Case Control Studies

- **Advantages:**
 - many exposures can be investigated in same study
 - specific exposures can be measured in detail (job exposure matrix)
 - Can be nested within an industrial cohort (to aid in control selection)
 - useful for rare outcomes
- **Disadvantages:**
 - Proper control selection may be difficult
(i.e. sometimes difficult to specify the source population that gave rise to the cases)
 - recall bias in exposure determination

Case Control Studies

- **Relevance to occupational health:**
 - frequently used design, especially in cancer epidemiology
 - when done properly, case control studies yield results as valid as those obtained from cohort studies

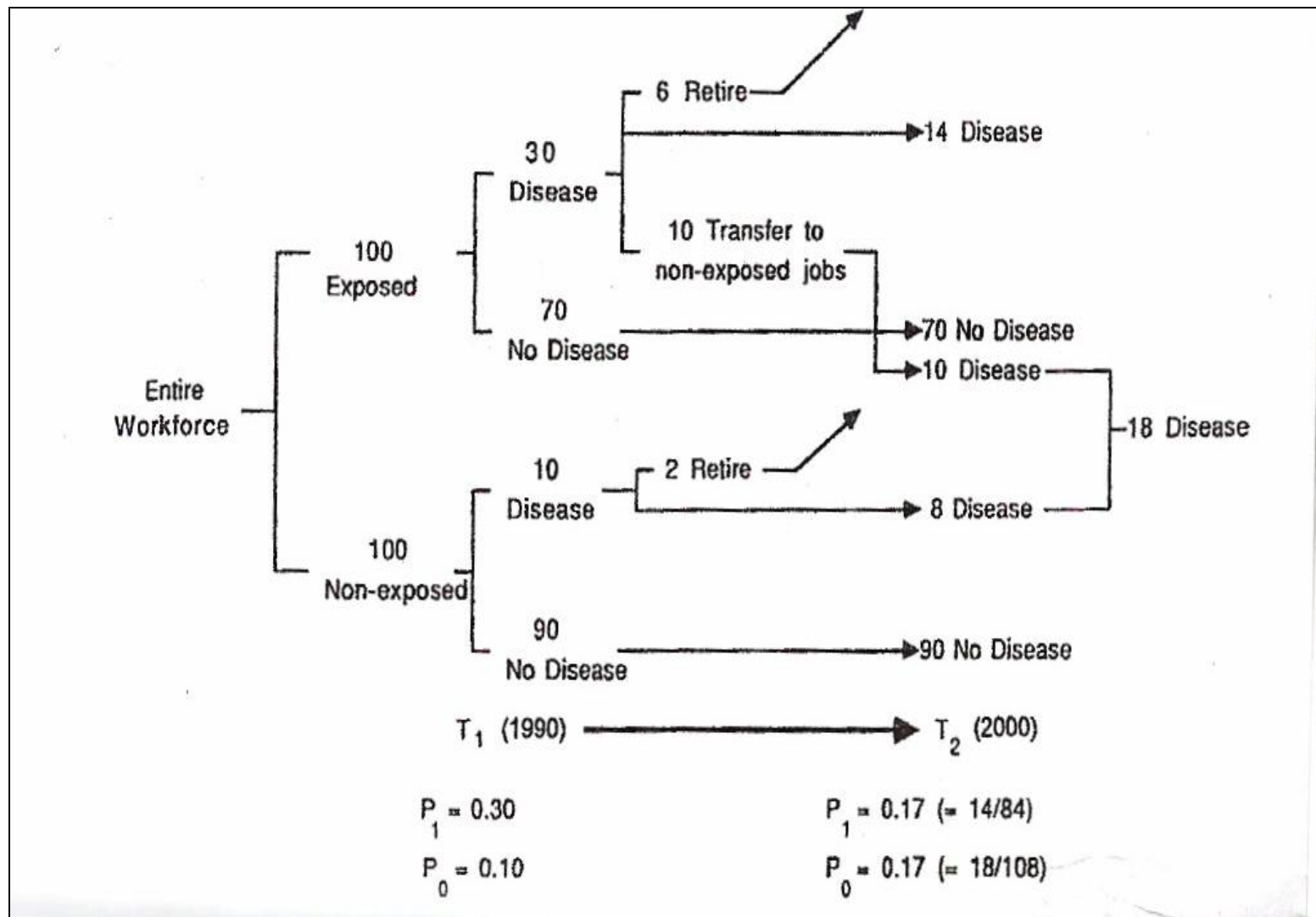
Cross Sectional Studies

- **Main feature:**
 - outcome and exposure determined at same time
- **Advantages:**
 - relatively easy to carry out
 - less expensive than prospective studies
 - permits study of outcomes for which data would not be collected on a routine basis

Cross Sectional Studies

- **Disadvantages:**
 - problem of temporality:
 - can't ensure that exposure preceded outcome
 - if studying currently employed workers, potential for selection bias exists
 - prevalent cases:
 - study may identify factors associated with duration rather than incidence
 - recall bias for self-reported exposures

Example of Selection Bias in Cross Sectional Study



Cross Sectional Studies

- **Relevance to occupational health:**
 - used extensively in studies of nonfatal diseases or physiologic responses to workplace exposures (morbidity studies)

Examples:

- Noise induced hearing loss
- Hand-arm vibration syndrome
- Toxic effects from metals, solvents

Causation and Meta-Analysis

Determination of Causation (Modified Bradford Hill)

- **Critical appraisal of studies addressing the specific causal question**
 - The importance of evaluating study design and sources of bias
- **Specific factors to be considered**
 - Experimental Evidence
 - Strength
 - Consistency
 - Temporality
 - Dose -Response (Biologic Gradient)
 - Epidemiologic Sense
 - Biologic Sense (Plausibility)
 - Specificity
 - Analogy

Evolution of Causation Theory

- Koch's Postulates 1882- emphasized **specificity** in evaluation of infectious disease causation
- Bradford Hill 1965 – factors to be considered (“viewpoints”) for chronic disease causation
- Later there was more emphasis on study quality and sources of bias
- Meta-analysis - emphasized the precision of the estimate of the **strength** of the association

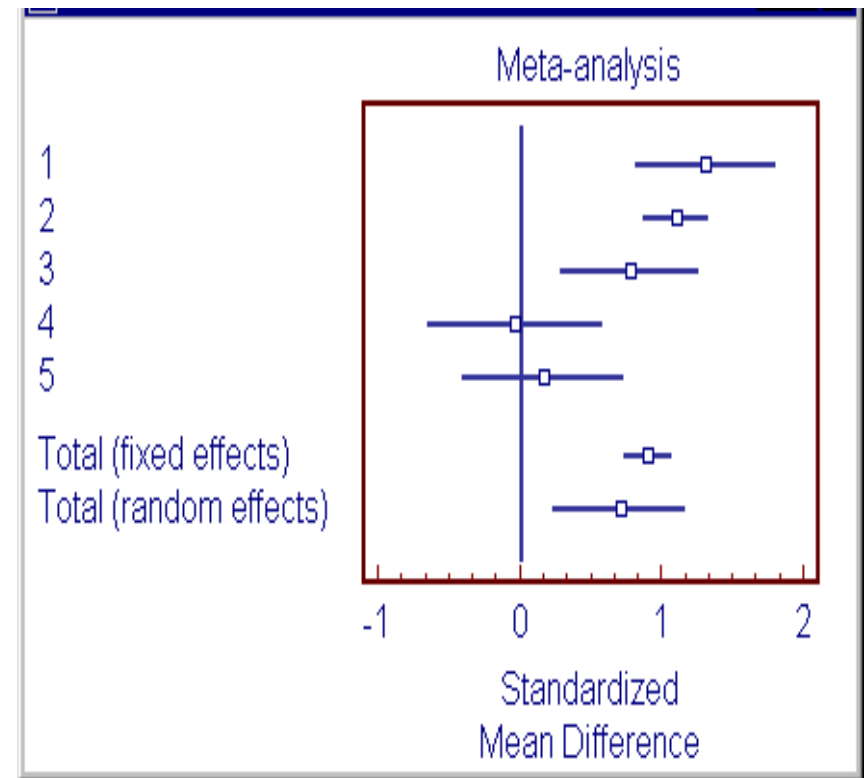
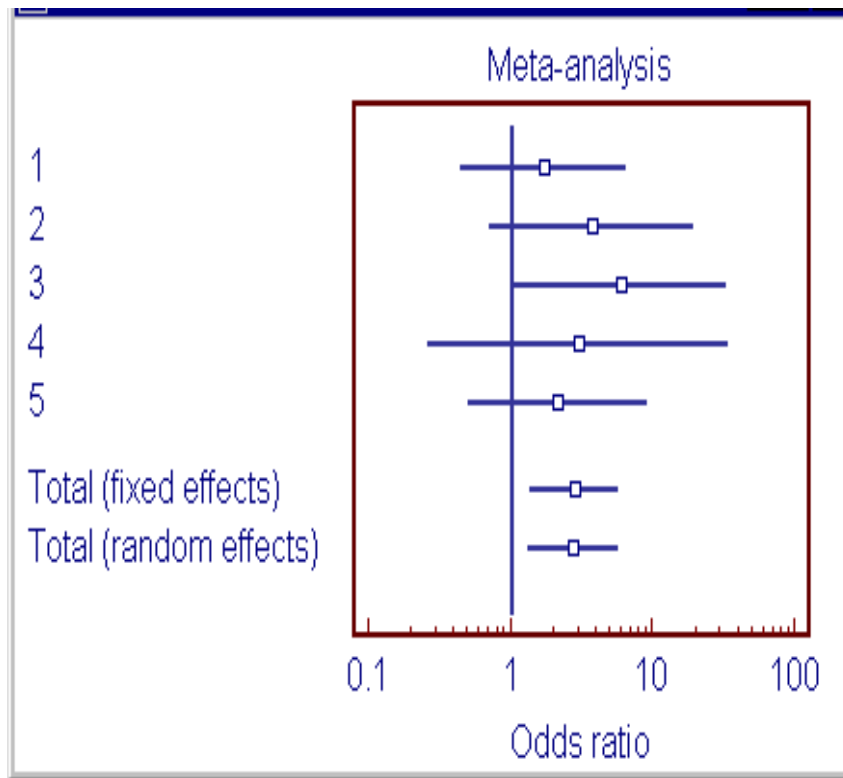
Types of Epidemiological Literature Review

- Meta-analysis
 - Method that critically reviews and combines the results of previous research
- Other types of review:
 - Systematic review
 - Narrative review

Meta-Analysis

- **Components:**
 - **Qualitative:**
 - systematic review of quality of component studies
 - **Quantitative:**
 - combine smaller studies to obtain increased power to detect an effect due to exposure/treatment

Graphical Display of Individual Study Results



Meta-Analysis: Overview of Process

- Formulate objectives (study question) and protocol
- Literature search
- Select studies for review
 - Based on inclusion/exclusion criteria
- Assess characteristics and quality of individual studies
- Determine which studies should be pooled
- Evaluation of results
 - Examine heterogeneity of individual studies
 - Combine studies and estimate overall effect size
- Sensitivity analysis
 - Vary the assumptions and see how this affects the conclusions
- Conclusions/recommendations

Overall Role of Meta-Analysis in Occupational Health

- Still some debate about when results of meta-analysis are sufficient to resolve controversy or when a definitive large study is needed
 - Meta-analysis should be treated as an aid to comparisons of studies and not focused solely on an estimation of an overall effect size
- Even if new observational study is needed meta-analysis of previous studies may be helpful for planning future research
 - Estimation of possible effect size
 - Identification of key sources of heterogeneity (confounders, effect modifiers)

Application of Epidemiological Studies to Determine Individual Probability of Causation

- Etiological fraction among the exposed (EFE)
- $$EFE = \frac{RR - 1}{RR}$$

= proportion of outcomes in the exposed group that is due to the exposure
- Used to determine the probability that a worker with the outcome of interest and a specific exposure developed that outcome due to the exposure

Etiologic Fraction Among the Exposed

If $RR = 2$

$$\begin{aligned} EFE &= (RR-1 / RR) \\ &= (2-1) / 2 \\ &= 0.5 \end{aligned}$$

If the $RR= 2$ the probability of causation in an exposed individual with the outcome is 50%