

# Systematic Reviews of Non-randomized Studies

---

*Workshop Day 2 PM  
November 20, 2007*



*George A Wells  
University of Ottawa*

## Outline

---

Introduction to Observational Study Designs

Assessing Quality of Non-randomized Studies for  
Systematic Reviews

Steps in a Systematic Review of Non-randomized  
Studies

## *Introduction to Observational Study Designs*

# **Observational Studies**

---

**Experimental studies:** investigators allocate subjects to different 'groups' and collect and analyze data on the subjects

**Observational studies:** investigators collect and analyze data on subjects as they naturally divide themselves by potentially significant variables into different 'groups'

## *Introduction to Observational Study Designs*

# **Overview**

---

- Making Comparisons
- Experimental Studies
  - Randomized controlled trials
- Observational studies
  - Cohort studies
  - Case-control studies
  - Cross-sectional studies
- Assessing the Quality of Observational Studies
- Assembling the Evidence

*Introduction to Observational Study Designs*  
**Case Study**

---

Does treatment of postmenopausal women with hormone replacement therapy (HRT) **cause** connective tissue disease (CHD)?

HRT recipients have subsequently developed clinically significant symptoms of CHD

*Introduction to Observational Study Designs*  
**Case Study**

---

To assess whether existing studies provide scientific evidence of an association between HRT and CHD

To do so, a systematic review of published studies was performed


**Studying whether one thing CAUSES another is a challenging task!**

## Introduction to Observational Study Designs


# Case Study

---

### First: What is the right question?

 This is not the right question! CHD occurs in women not on HRT and so the relative frequency is of importance.

What is the frequency of occurrence of CHD in women on HRT?

 This is the right question! Making a comparison

Were women on HRT more likely to develop CHD than women not on HRT?

## Introduction to Observational Study Designs

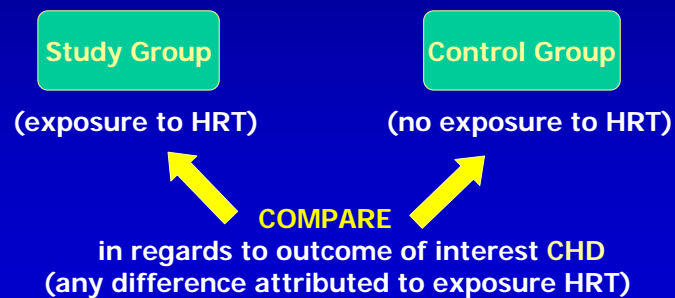
# Making Comparisons

---

To answer the question, we need to...

Establish a standard of comparison

Conduct a comparative study



## *Introduction to Observational Study Designs*

# **Case Study**

---

### **Second: What is the right study design?**

The best scientific test of the relationship between HRT and CHD would be a RCT in which women, initially free of CHD who consent to the trial are randomly allocated to receive or not receive HRT, and then followed to assess the incidence of CHD

#### **Evidence from other study designs**

If RCT not possible, other types of evidence are needed, although none by themselves will be as compelling

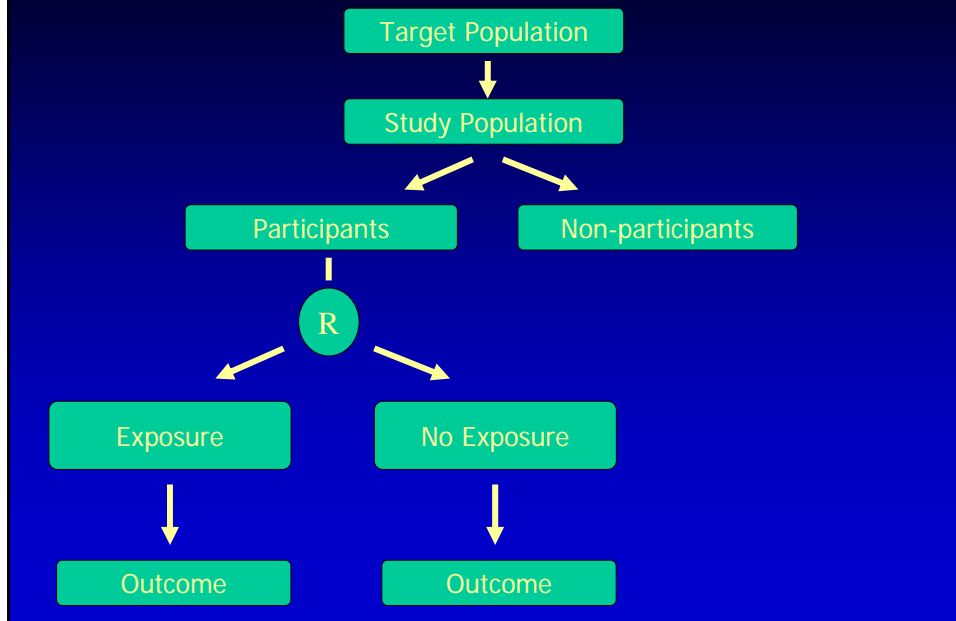
## *Introduction to Observational Study Designs*

# **Overview**

---

- **Making Comparisons**
- **Experimental Studies**
  - Randomized controlled trials
- **Observational studies**
  - Cohort studies
  - Case-control studies
  - Cross-sectional studies
- **Assessing the Quality of Observational Studies**
- **Assembling the Evidence**

## Structure: Randomized Controlled Trial



## Introduction to Observational Study Designs Randomized controlled trials

### Data:

		Outcome	
		Yes	No
R	Exposed	a	b
	Not exposed	c	d

## *Introduction to Observational Study Designs*

### ***Fundamental fact***

---

A large, well-designed, properly conducted randomized trial



yields comparison groups similar in every way except for the exposure



attribute differences in outcome between groups to differences in exposure

## *Introduction to Observational Study Designs*

### ***Case Study***

---

Women's Health Initiative (WHI) evaluated exposure (HRT) and outcome (CHD)

RCTs (HRT vs control) with CHD considered as adverse event

*Introduction to Observational Study Designs*  
***Randomized controlled trials***

---

**Potential Strengths**

can demonstrate causal relationships with a high degree of confidence

can control 'exposure' levels

*Introduction to Observational Study Designs*  
***Randomized controlled trials***

---

**Potential Weaknesses**

limited applicability due to ethic considerations and artificial experimental setting

differential rates of noncompliance, withdrawals and losses to follow-up

costly and logistically time consuming

sample size due to recruitment (volunteers, eligibility)

external validity

## *Introduction to Observational Study Designs*

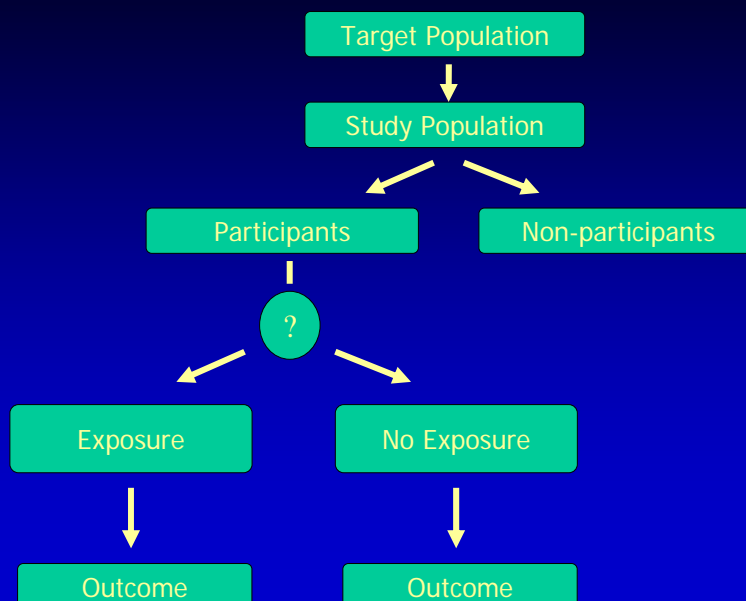
# **Overview**

---

- Making Comparisons
- Experimental Studies
  - Randomized controlled trials
- Observational studies
  - Cohort studies
  - Case-control studies
  - Cross-sectional studies
- Assessing the Quality of Observational Studies
- Assembling the Evidence

### *Structure: Cohort Study*

---



## Introduction to Observational Study Designs

# Cohort Study

---

### Distinguishing Features:

Groups of persons (cohorts) to be studied are defined in terms of characteristics that manifest prior to outcome under study (exposed and not exposed)

Groups observed over period of time to determine frequency of outcome among them

### Then

rate for outcome for those exposed is compared to the corresponding rate for those not exposed

### If

rates are significantly different then an association is said to exist between the exposure and outcome


## Introduction to Observational Study Designs

# Cohort Study

---

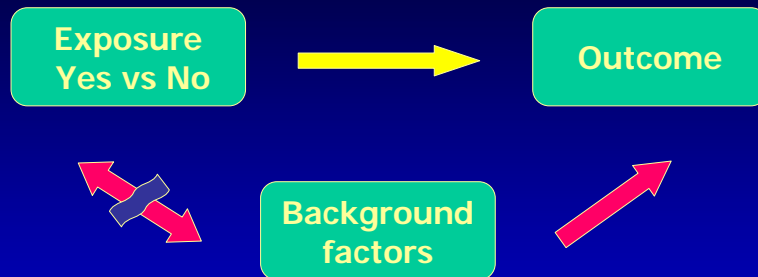
### Data:

		Outcome	
		Yes	No
Exposure	Exposed	a	b
	Not exposed	c	d



## Introduction to Observational Study Designs

# Making Comparisons



1. Design
  - match
  - stratified sampling
  - restrict inclusion
2. Analysis
  - standardize
  - stratify (subgroups)
  - adjust (statistical model)
3. Randomize

## Introduction to Quantitative Research Designs

# Cohort Studies

### Types of Cohort Studies:

**Prospective cohort studies:** classifying subjects without the outcome of interest on the basis of exposure status; follow cohort longitudinally into the future to determine outcome

**Retrospective cohort studies:** historical cohort reconstructed from existing data sources at the beginning of the study; subjects classified by exposure status; determine outcome status (possible follow-up)

## *Introduction to Observational Study Designs*

# **Case Study**

---

**14 cohort studies** were found evaluating exposure (HRT) and outcome (CHD)

## *Introduction to Observational Study Designs*

# **Cohort Studies**

---

### Potential Strengths

demonstrate appropriate temporal sequence between exposure and outcome

permit direct calculation of incidence rates in exposed and unexposed groups

allow multiple outcomes to be evaluated

provide indication of incubation or latency period of outcome

suitable for studying rare exposures

outcome status determination unlikely to bias exposure status determination

can help establish cause-effect relationships

## *Introduction to Observational Study Designs*

# ***Cohort Studies***

---

### Potential Weaknesses

complex and costly (large numbers and long follow-up needed to study outcomes with low incidence and long latency periods)

loss to follow-up bias

exposure misclassification due to changes in exposure during follow-up period or because of inadequate information in retrospective designs

outcome misclassifications when advances in detection during follow-up phase question earlier classifications or knowledge of exposure status biases assessment of outcome status

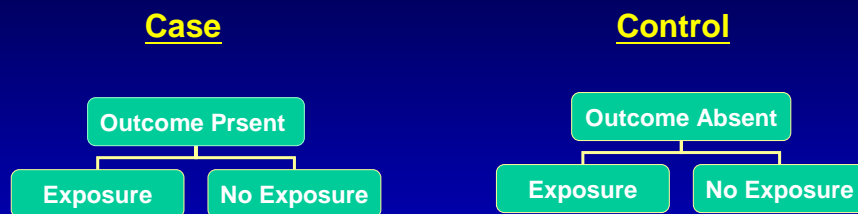
## *Introduction to Observational Study Designs*

# ***Overview***

---

- Making Comparisons
- Experimental Studies
  - Randomized controlled trials
- Observational studies
  - Cohort studies
  - Case-control studies
  - Cross-sectional studies
- Assessing the Quality of Observational Studies
- Assembling the Evidence

## Structure: of Case-Control Study



## Introduction to Observational Study Designs Case-Control Study

### Distinguishing Features:

Groups of patients assembled in terms of whether they do (cases) or do not (controls) have the outcome under study

Groups are then compared with respect the exposure under study judged to be of possible relevance to occurrence of outcome

#### **Then**

number of cases exposed or not exposed is compared to corresponding number of controls

#### **If**

significantly more cases than controls exposed then an association between the outcome and exposure is inferred

## Introduction to Observational Study Designs

# Case-Control Study

---

Data:

Exposure	Outcome	
	Yes	No
Exposed	a	b
Not exposed	c	d



## Introduction to Observational Study Designs

# Case Study

---

16 case-control studies were found evaluating exposure (HRT) and outcome (CHD)

## *Introduction to Observational Study Designs*

# ***Case-Control Studies***

---

### Potential Strengths

- quick and inexpensive
- appropriate for studying rare outcomes
- can conduct with moderate number of subjects
- allows multiple potential risk factors to be examined
- can help in establishing cause-effect relationships

## *Introduction to Observational Study Designs*

# ***Case-Control Studies***

---

### Potential Weaknesses

- cannot determine incidence rates in exposed and unexposed subjects
- not appropriate for studying rare exposures
- availability of accurate information on prior exposure or confounders and with same quality in cases and controls
- difficulty identifying comparable case and control groups

## *Introduction to Observational Study Designs*

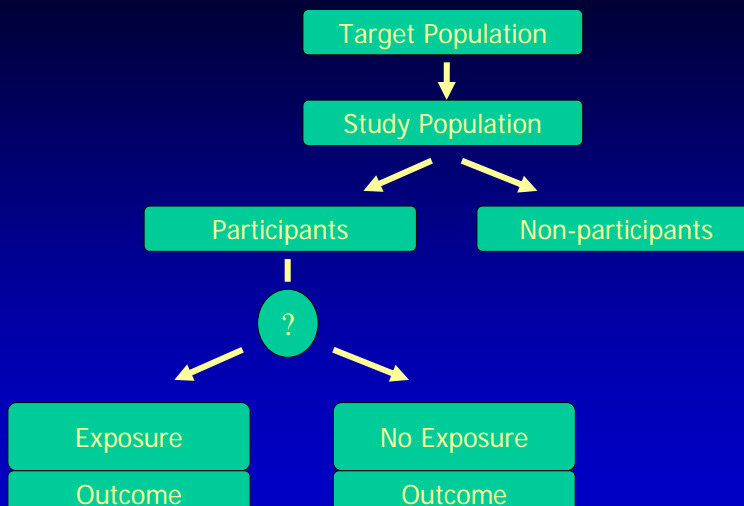
# **Overview**

---

- Making Comparisons
- Experimental Studies
  - Randomized controlled trials
- Observational studies
  - Cohort studies
  - Case-control studies
  - Cross-sectional studies
- Assessing the Quality of Observational Studies
- Assembling the Evidence

### *Structure: Cross-sectional Study*

---



## *Introduction to Observational Study Designs*

# ***Cross-sectional Study***

---

### Distinguishing Features:

Groups of persons to be studied are defined simultaneously in terms of exposure and outcome under study

### **Then**

number with outcome for those exposed is compared to the corresponding number for those not exposed

### **If**

numbers are significantly different then an association is said to exist between the exposure and outcome

## *Introduction to Observational Study Designs*

# ***Cross-sectional Study***

---

### Data:

	Outcome	
Exposure	Yes	No
Exposed	a	b
Not exposed	c	d



*Introduction to Observational Study Designs*  
**Case Study**

---

? **cross-sectional studies** were found evaluating exposure (HRT) and outcome (CHD)

*Introduction to Observational Study Designs*  
**Cross-sectional Studies**

---

Potential Strengths

quick and inexpensive

representativeness of findings to target population since probability sampling easier to implement

may be only appropriate design when incidence cannot be established

## *Introduction to Observational Study Designs*

# ***Cross-sectional Studies***

---

### Potential Weaknesses

correct temporal sequence between exposure and outcome usually cannot be determined

prevalence -incidence bias (transitory or fatal cases more likely to be missed)

unsuitable for study of rare outcomes

associations based on current exposure or outcome status may not be representative of past exposure or outcome status

## *Introduction to Observational Study Designs*

# ***Overview***

---

- Making Comparisons
- Experimental Studies
  - Randomized controlled trials
- Observational studies
  - Cohort studies
  - Case-control studies
  - Cross-sectional studies
- Assessing the Quality of Observational Studies
- Assembling the Evidence

## *Introduction to Observational Study Designs* *Assessing Quality of Observational Studies for Systematic Reviews*

---



### **Newcastle Ottawa Scale (NOS)**



focus on extent bias is minimized  
transparent and parsimonious system  
'quick and easy'



### ***Development: Grouping Items***

- **Case-Control studies**
  - Selection of case and controls
  - Comparability of cases and controls
  - Ascertainment of exposure
- **Cohort studies**
  - Selection of cohorts
  - Comparability of cohorts
  - Assessment of outcome



## **Newcastle-Ottawa Scale: Case-Control Studies**

**Selection (4 items)**

**Comparability (1 item)**

**Exposure (3 items)**

- Identify a 'high' quality choice for an item with a ★

### Selection

- 1. Is the case definition adequate?**
  - a) yes, with independent validation ★
  - b) yes, eg record linkage or based on self reports
  - c) no description
- 2. Representativeness of the cases**
  - a) consecutive or obviously representative series of cases ★
  - b) potential for selection biases or not stated
- 3. Selection of Controls**
  - a) community controls ★
  - b) hospital controls
  - c) no description
- 4. Definition of Controls**
  - a) no history of disease (endpoint) ★
  - b) no description of source

## Comparability

### 1. Comparability of cases and controls on the basis of the design or analysis

- a) study controls for \_\_\_\_\_ ★  
(select the most important factor)
- b) study controls for any additional factor ★  
(This criteria could be modified to indicate specific control for a second important factor.)

## Exposure

### 1. Ascertainment of exposure

- a) secure record (eg surgical records) ★
- b) structured interview where blind to case/control status ★
- c) interview not blinded to case/control status
- d) written self report or medical record only
- e) no description

### 2. Same method of ascertainment for cases and controls

- a) yes ★
- b) no

### 3. Non-Response Rate

- a) same rate for both groups ★
- b) non respondents described
- c) rate different and no designation



## Development: Grouping Items

- **Case-Control studies**

Selection of case and controls (4 items)



Comparability of cases and controls (1 item)



Ascertainment of exposure (3 items)



## **Newcastle-Ottawa Scale: Cohort Studies**

Selection (4 items)

Comparability (1 item)

Outcome (3 items)

- Identify a 'high' quality choice for an item with a 

## Selection

1. Representativeness of the exposed cohort
  - a) truly representative of the average \_\_\_\_\_ (describe) in the community ★
  - b) somewhat representative of the average \_\_\_\_\_ in the community ★
  - c) selected group of users eg nurses, volunteers
  - d) no description of the derivation of the cohort
2. Selection of the non exposed cohort
  - a) drawn from the same community as the exposed cohort ★
  - b) drawn from a different source
  - c) no description of the derivation of the non exposed cohort
3. Ascertainment of exposure
  - a) secure record (eg surgical records) ★
  - b) structured interview ★
  - c) written self report
  - d) no description
4. Demonstration that outcome of interest was not present at start of study
  - a) yes ★
  - b) no

## Comparability

1. Comparability of cohorts on the basis of the design or analysis
  - a) study controls for \_\_\_\_\_ ★  
(select the most important factor)
  - b) study controls for any additional factor ★  
(This criteria could be modified to indicate specific control for a second important factor.)

## Outcome

### 1. Assessment of outcome

- a) independent blind assessment ★
- b) record linkage ★
- c) self report
- d) no description

### 2. Was follow up long enough for outcomes to occur

- a) yes (select an adequate follow up period for outcome of interest) ★
- b) no

### 3. Adequacy of follow up of cohorts

- a) complete follow up - all subjects accounted for ★
- b) subjects lost to follow up unlikely to introduce bias - small number lost - > \_\_\_ % (select an adequate %) follow up, or description of those lost) ★
- c) follow up rate < \_\_\_ % (select an adequate %) and no description of those lost
- d) no statement



## Development: Grouping Items

### • **Case-Control studies**

- Selection of case and controls (4 items) ★ ★ ★ ★
- Comparability of cases and controls (1 item) ★ ★
- Ascertainment of exposure (3 items) ★ ★ ★

### • **Cohort studies**

- Selection of cohorts (4 items) ★ ★ ★ ★
- Comparability of cohorts (1 item) ★ ★
- Assessment of outcome (3 items) ★ ★ ★

**NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE  
CASE CONTROL STUDIES**

**Note:** A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

**Selection**

- 1) Is the case definition adequate?
  - a) yes, with independent validation
  - b) yes, eg record linkage or based on self reports
  - c) no description
- 2) Representativeness of the cases
  - a) consecutive or obviously representative series of cases
  - b) potential for selection biases or not stated
- 3) Selection of Controls
  - a) community controls
  - b) hospital controls
  - c) no description
- 4) Definition of Controls
  - a) no history of disease (endpoint)
  - b) no description of source

**Comparability**

- 1) Comparability of cases and controls on the basis of the design or analysis
  - a) study controls for \_\_\_\_\_ (Select the most important factor.)
  - b) study controls for any additional factor  (This criteria could be modified to indicate specific control for a second important factor.)

**Exposure**

- 1) Ascertainment of exposure
  - a) secure record (eg surgical records)
  - b) structured interview where blind to case/control status
  - c) interview not blinded to case/control status
  - d) written self report or medical record only
  - e) no description
- 2) Same method of ascertainment for cases and controls
  - a) yes
  - b) no
- 3) Non-Response rate
  - a) same rate for both groups
  - b) non respondents described
  - c) rate different and no designation

**Case-Control Studies: NOS Quality Assessment**

	<b>Selection</b>	<b>Comparability</b>	<b>Exposure</b>
Adam 1981	☆☆☆☆	☆	☆☆
Beard 1989	☆☆☆☆	☆☆	☆☆
Croft 1989	☆☆☆☆	☆☆	☆☆
Grodstein 1997	☆☆	☆☆	☆☆
Heckbert 1997	☆☆☆☆	☆☆	☆☆☆
LaVecchia 1987	☆☆	☆☆	☆☆
Mann 1994	☆☆☆	☆☆	☆☆
Pfeffer 1978	☆☆☆☆	☆☆	☆☆☆
Rosenberg 1976	☆	☆☆	☆
Rosenberg 1980	☆☆☆	☆☆	☆☆
Rosenberg 1993	☆☆☆☆	☆☆	☆☆
Ross 1981	☆☆☆☆	☆☆	☆☆
Sidney 1997	☆☆☆☆	☆☆	☆☆
Szklo 1984	☆☆☆	☆☆	☆☆
Talbott 1977	☆☆☆	☆☆	☆
Thompson 1989	☆☆☆	☆☆	☆

**NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE  
COHORT STUDIES**

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

**Selection**

- 1) Representativeness of the exposed cohort  
 a) truly representative of the average \_\_\_\_\_ (describe) in the community   
 b) somewhat representative of the average \_\_\_\_\_ in the community   
 c) selected group of users eg nurses, volunteers  
 d) no description of the derivation of the cohort
- 2) Selection of the non exposed cohort  
 a) drawn from the same community as the exposed cohort   
 b) drawn from a different source  
 c) no description of the derivation of the non exposed cohort
- 3) Ascertainment of exposure  
 a) secure record (eg surgical records)   
 b) structured interview   
 c) written self report  
 d) no description
- 4) Demonstration that outcome of interest was not present at start of study  
 a) yes   
 b) no

**Comparability**

- 1) Comparability of cohorts on the basis of the design or analysis  
 a) study controls for \_\_\_\_\_ (select the most important factor)   
 b) study controls for any additional factor  (This criteria could be modified to indicate specific control for a second important factor.)

**Outcome**

- 1) Assessment of outcome  
 a) independent blind assessment   
 b) record linkage   
 c) self report  
 d) no description
- 2) Was follow-up long enough for outcomes to occur  
 a) yes (select an adequate follow up period for outcome of interest)   
 b) no
- 3) Adequacy of follow up of cohorts  
 a) complete follow up - all subjects accounted for   
 b) subjects lost to follow up unlikely to introduce bias - small number lost - > \_\_\_\_ % (select an adequate % follow up, or description provided of those lost)   
 c) follow up rate < \_\_\_\_% (select an adequate %) and no description of those lost  
 d) no statement

**Cohort Studies: NOS Quality Assessment**

	Selection	Comparability	Outcome
Avila 1990	★★★★	★★	★★★
Bush 1987	★★	★★	★★★
Cauley 1997	★★★	★★	★★★
Criqui 1998	★★	★★	★★★
Ettinger 1996	★★★★	★★	★★★
Folsom 1995	★★★	★★	★★
Grodstein 1996	★★	★★	★★★
Henderson 1991	★★★	★★	★★★
Lafferty 1994	★★★★	★	★★★
Lauritzen 183	★★★★	★★	★
Petitti 1987	★★	★★	★★
Sourander 1998	★★	★★	★★
Wilson 1985	★★★★	★★	★★★
Wolf 1996	★★★	★★	★★★

## Introduction to Observational Study Designs

# Overview

---

- Making Comparisons
- Experimental Studies
  - Randomized controlled trials
- Observational studies
  - Cohort studies
  - Case-control studies
  - Cross-sectional studies
- Assessing the Quality of Observational Studies
- Assembling the Evidence

## Long Term Hormone Replacement Therapy and Coronary Heart Disease

---



### Steps of a Cochrane Systematic Review

- Clearly formulated question

Is there a relationship between hormone replacement therapy and the incidence of coronary heart disease in postmenopausal women

**Definition of the population:** postmenopausal women in replacement therapy  
**Definition of the outcome:** fatal (cardiovascular) and nonfatal (myocardial infarction) angina, and sudden cardiac death route over a year (ever, current, past)

## **Long Term Hormone Replacement Therapy and Coronary Heart Disease**



### **Steps of a Cochrane Systematic Review**

- Comprehensive data search

#### **Electronic Search of:**

MEDLINE (1966 to May 2000)  
Current Contents (to May 2000)

#### **Other Data Sources:**

review of references cited in retrieved articles

## **Long Term Hormone Replacement Therapy and Coronary Heart Disease**



### **Steps of a Cochrane Systematic Review**

- Unbiased selection and extraction process

#### **Selection:**

2 independent reviewers selected studies using detailed criteria (consensus)

#### **Extraction:**

2 independent reviewers extracted data using pre-determined forms: study design, population characteristics, exposure to hormone replacement therapy, outcomes measures, results (consensus)

# Long Term Hormone Replacement Therapy and Coronary Heart Disease



## Steps of a Cochrane Systematic Review

- Critical appraisal of data

### Table of Study Characteristics

- Size of study
- Characteristics of study patients
- Details of specific interventions used
- Details of outcomes assessed

### Methodologic quality assessment

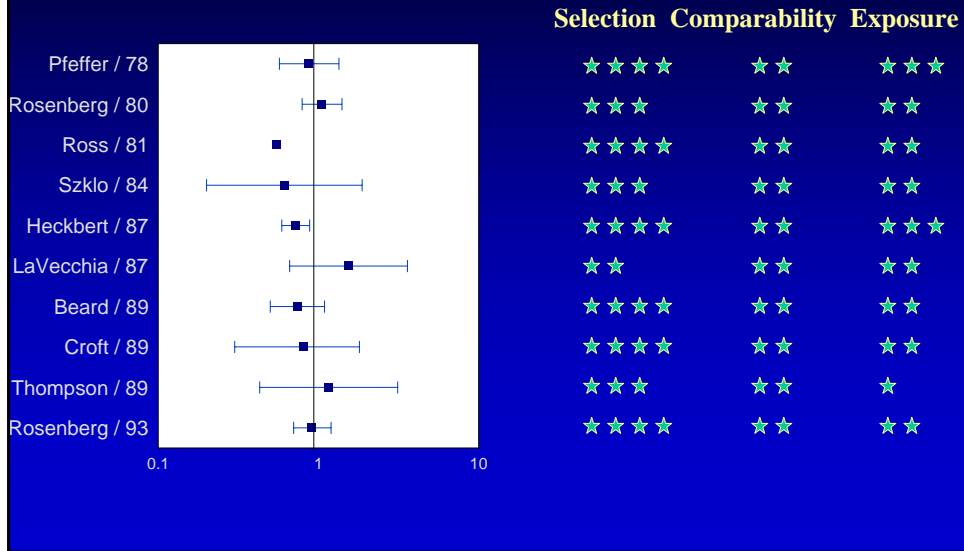
## Results

- 16 case-control or cross-sectional studies
- 14 cohort studies

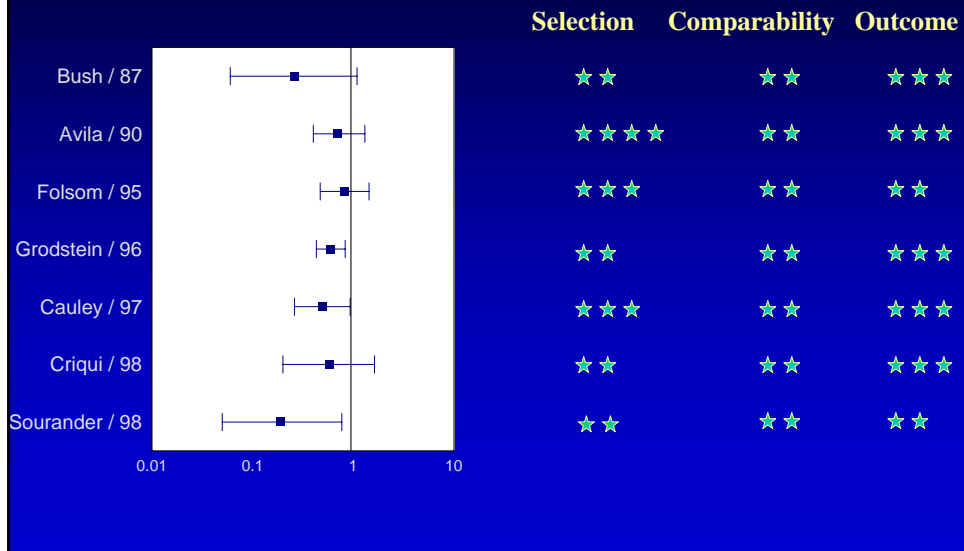
### Quantification of Effects:

- Exposure (ever, current, past)
- Outcome (fatal, non-fatal, both)
- Effect estimates (EE)
  - Relative Risk (RR)
  - Odds Ratio (OR)
- Adjusted effect estimates
- Effects vs population, follow-up periods, quality etc

**Adjusted Effect Estimates for:  
Coronary Heart Disease (All Events)  
Estrogen (Ever Use)  
Case-Control Studies**



**Adjusted Effect Estimates for:  
Coronary Heart Disease (All Events)  
Estrogen (Current Use)  
Cohort Studies**



## *Introduction to Observational Study Designs*

# **Case Study**

---

Strength of Association:

In 14 cohort, 16 case-control / cross-sectional studies, no association was found between HRT and CHD

## *Limitations of Randomized Trials*

# **Some starting points**

---

### Limitations of the RCT\*

RCTs are expensive and are hence **limited in size** and/or **length of follow-up**. They are therefore usually aimed at estimating clinical efficacy. For this reason, the RCT is usually a poor instrument for estimating **rare adverse effects or delayed or cumulative effects**. Also, subjects included in RCTs are often **not representative of the general patient population**. For example, the very young, the very old, the very frail and patients on multiple therapies are often excluded from trials.

\* Hierarchy of evidence: data from different trials by Alain Li Wan Po PhD, DipEcon, FPSNI, FRPharmS, FRSC, FRSS

## *Limitations of Randomized Trials*

### **Some starting points**

---

#### Reasons for doing a SR of NRS of effectiveness \*

(One of 3 reasons identified)

To provide evidence of effects (benefit or harm) that cannot be adequately studied in RCTs, such as **long-term** and **rare outcomes**, or effects in patient **groups not recruited** to RCTs (such as children, pregnant women, the elderly).

\* Handbook: Nonrandomised Study Methods Group, Cochrane Collaboration (in preparation)

## *Limitations of Randomized Trials*

### **Issues to consider**

---

- Choice of primary efficacy outcome
- Choice of duration of trial
- Inclusion of complicated patients
- Detection of adverse events
- Guidance on 'varying' the intervention

## *Limitations of Randomized Trials*

# **Adverse events**

---

- low power to detect important differences in
  - adverse event rates
  - adverse event profiles
  - new but rare adverse events

## *Limitations of Randomized Trials*

# **Adverse events**

---

- Observational studies
  - Analytic studies
    - Case-control studies
    - Cohort studies

## Limitations of Randomized Trials

# Adverse events

---

### Case-Control Study

Primary group classification is by outcome ('event') and the secondary by treatment



## Limitations of Randomized Trials

# Adverse events

---

### Cohort Study

Primary group classification is by treatment and the secondary by outcome ('event')



# **Systematic Reviews of Non-randomized Studies**

---

*Workshop Day 2 PM  
November 20, 2007*