

# **Analytic Study Design**

**DMID/ICSSC**

**Dar es Salaam, Tanzania**

**January 17-22, 2010**

**Kenneth F. Schulz, PhD**



© 2001 Harvey Nelson



© 2002 Alfredo L. Fort

# Cohort Studies



© 2002 Kelli M. Donley



© 1997 Ivin Morgan/CDC

# Example of a Prospective Cohort Design: Treatment of Severe Malaria in Children

## The Present



### Sample

PTX present

PTX absent

## The Future

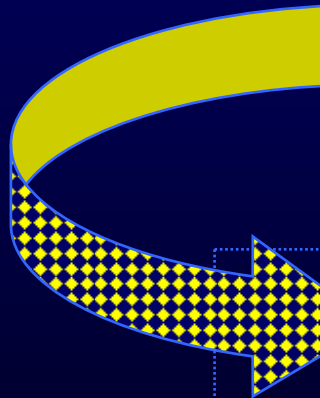
Death

No Death

Death

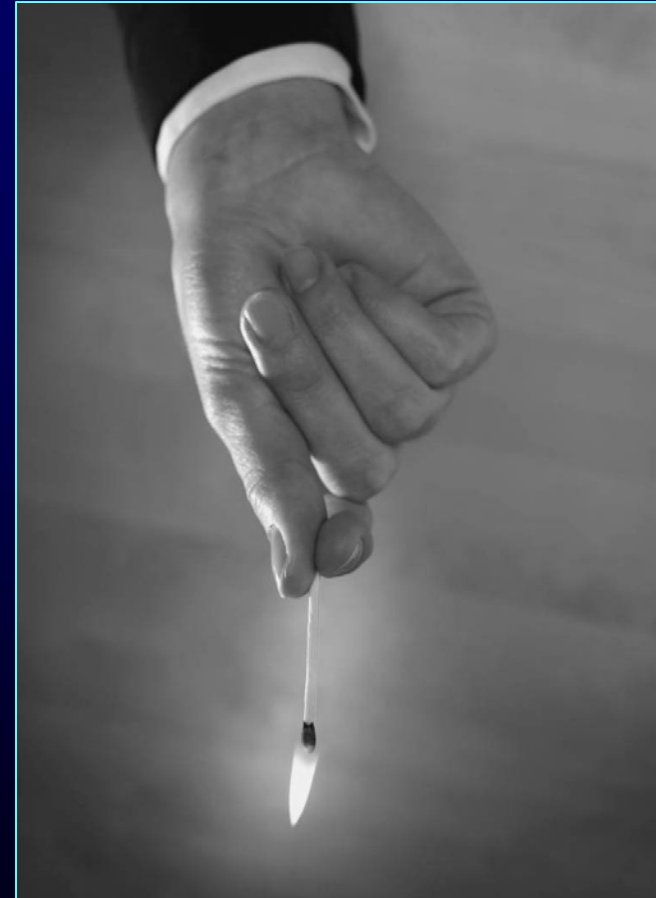
No Death

Population



# In a Cohort Study

- ❖ The outcome has not occurred when the study starts . . . usually
- ❖ **Are retrospective cohort studies possible?**



# Bias: Cohort Compared to Case-Control

- ❖ **Prospective cohort studies do not have two potent causes of bias in the case-control studies**
  - ◆ **Selection bias in the control group**
  - ◆ **Ascertainment bias in measuring exposure**
- ❖ **While biased, tend to be less biased than case-control**

# **Additional Advantages of Cohort Studies**

---

**Because potential causative factors are measured before the outcome occurs, a cohort study can establish that they preceded outcome**

# Steps in Prospective Cohort Studies

1. Select a sample from the population
2. Measure exposures (predictor variables)
3. Measure confounding factors
4. Follow-up the cohort
5. Measure outcome variables
6. Analyze results

# Simple Cohort Study Analysis

	Exposed	Unexposed
Participants developing outcome	a	b
Participants not developing outcome	c	d
Total number of participants	$N_1$	$N_0$

$$\text{Relative risk RR} = \frac{\text{incidence of outcome in exposed group}}{\text{incidence of outcome in unexposed group}} = \frac{a/N_1}{b/N_0}$$

$$\text{Attributable risk} = \text{incidence of the outcome that can be attributed to the exposure} = a/N_1 - b/N_0$$

# Cohort Study Analysis Example

	Exposed	Unexposed
Outcome	5	5
No Outcome	95	195
Total	100	200

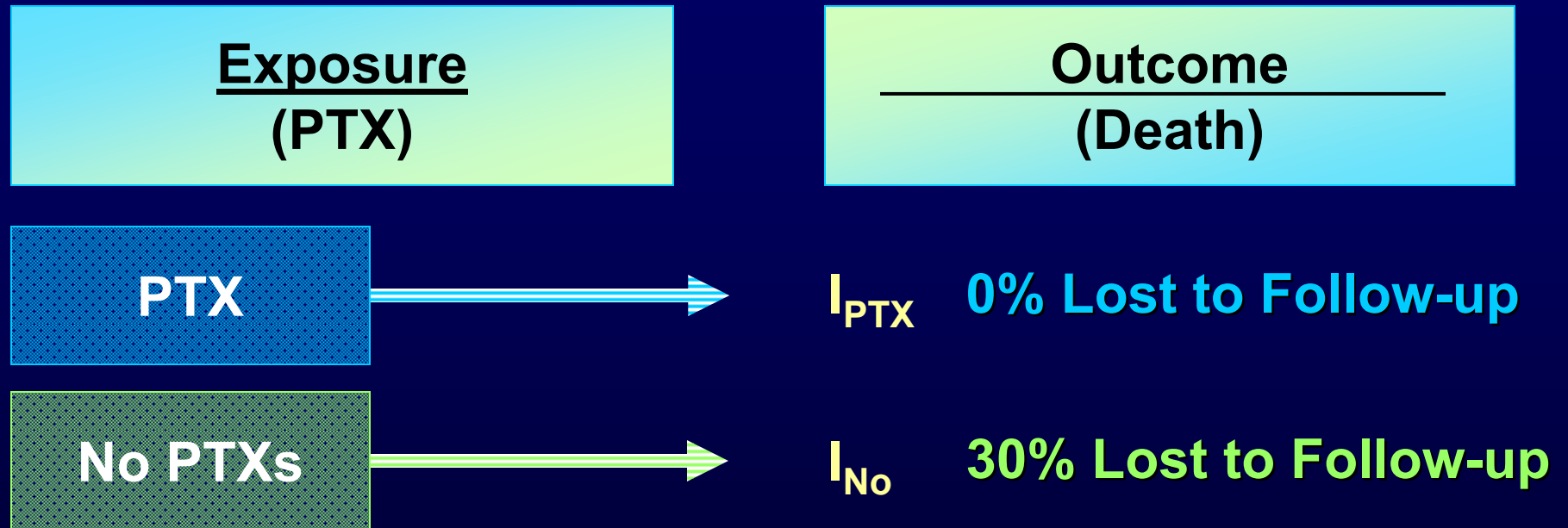
$$I_E = 5.0\%$$

$$I_{\bar{E}} = 2.5\%$$

$$RR = 2.0$$

$$AR = 2.5\%$$

# The Problem of Lost To Follow-up



- (Lost to follow-up have higher incidence of death)
- Measured death rate in No PTX group too low
- Biases comparison

--- Is the incidence rate in No PTX group lower than in the PTX group?

# The Problem of Lost To Follow-up

Exposure  
(PTX)

Outcome  
(Death)

PTX



$I_{PTX}$  20% Lost to Follow-up

No PTXs



$I_{No}$  20% Lost to Follow-up

- Lost to follow-ups have unknown impact on outcome
- Measured death rate in both groups too low or high
- Bias?

In reality, LFUs could be different in both groups

# Minimize Losses to Follow-up

- ❖ **Must** minimize losses
- ❖ Address procedures for minimizing losses in greater detail later in the week
  - ◆ Retention
  - ◆ Cohort and RCTs face similar issues

# **Ascertainment Bias (*Information Bias*) Potential**

**Since exposure is likely known by both the participants and study investigators, danger of diagnosis of outcome being influenced by exposure class**

# Minimize Ascertainment Bias in Cohort Studies - Determination of Outcomes

- ❖ **Consistent**
- ❖ **Equal for all exposure groups**
- ❖ **Establish explicit, objective criteria**
- ❖ **Outcomes should be assessed blindly, if possible**

# CONFOUNDING

		PTX	$\overline{\text{PTX}}$
Death	Yes	15	45
	No	85	55
		100	100

$$\text{RR} = \frac{15\%}{45\%} = 0.33$$

# Confounding Bias Occurs

- ❖ Not hypothetical
- ❖ Occurs in both cohort and case-control
- ❖ Confounding is confusing and needs convincing
  - ◆ Play with numbers
  - ◆ See page 37

In this example you suspect:

**SES is strongly associated with both PTX and death,**

**i.e., SES is a Confounder**

High SES

Death  
Yes  
No

	PTX	$\overline{\text{PTX}}$
Yes	9	3
No	81	27
	90	30

$$RR = \frac{10\%}{10\%} = 1.0$$

Low SES

Death  
Yes  
No

	PTX	$\overline{\text{PTX}}$
Yes	6	42
No	4	28
	10	70

$$RR = \frac{60\%}{60\%} = 1.0$$

# Example of Confounding

## Example of confounding in a hypothetical cohort study of PTX and death

- ◆ When the relative risk is controlled for the confounding effect of SES, the decreased risk (protective effect) disappears
- ◆ Play with the numbers in the salpingitis example on page 37

# Advantages of Cohort Studies

- ❖ Efficient with higher incidence (approx.  $> 20\%$ )
- ❖ Excellent for studying rare exposures
- ❖ Less opportunity for selection bias and ascertainment bias than case-control studies
- ❖ Clear temporal sequence of exposure and outcome
- ❖ Obtain incidence rates and relative risk
- ❖ Yields more understandable information than c-c

# Disadvantages of Cohort Studies

- ❖ Contains selection bias and probably more ascertainment bias than an RCT
- ❖ With rare outcomes, large sample sizes and relatively expensive to conduct
- ❖ Long-term follow-up difficult when the latency period for the outcome is long
- ❖ Follow-up may be difficult -- losses affect results
- ❖ Exposure status may change during study

## **At a minimum for a cohort study, address in the protocol:**

- ❖ **Entry criteria**
- ❖ **Definitions of the exposure groups (comparison groups) [and implications for selection bias]**
- ❖ **The planned procedures to achieve retention and follow-up of participants**
- ❖ **Selection and measurement of potential confounding factors**
- ❖ **Endpoint: Next sessions (Methods to ascertain outcomes, including blinding procedures, if any)**



# Case-Control Studies



# Incidence and Relative Risk From a Case-Control study

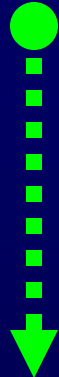
	E	$\bar{E}$	
Cases	10	90	100
Controls	90	110	200
	100	200	

Incidence rate in E ?

Incidence rate in  $\bar{E}$  ?

RR?


# Cohort Study

	Exposed	Unexposed	
Cases	a	b	
Controls	c	d	
	a + c	b + d	

Start with a + c and b + d; determine OUTCOME

# Case-Control Study

	Exposed	Unexposed	
Cases	a	b	a + b
Controls	c	d	c + d

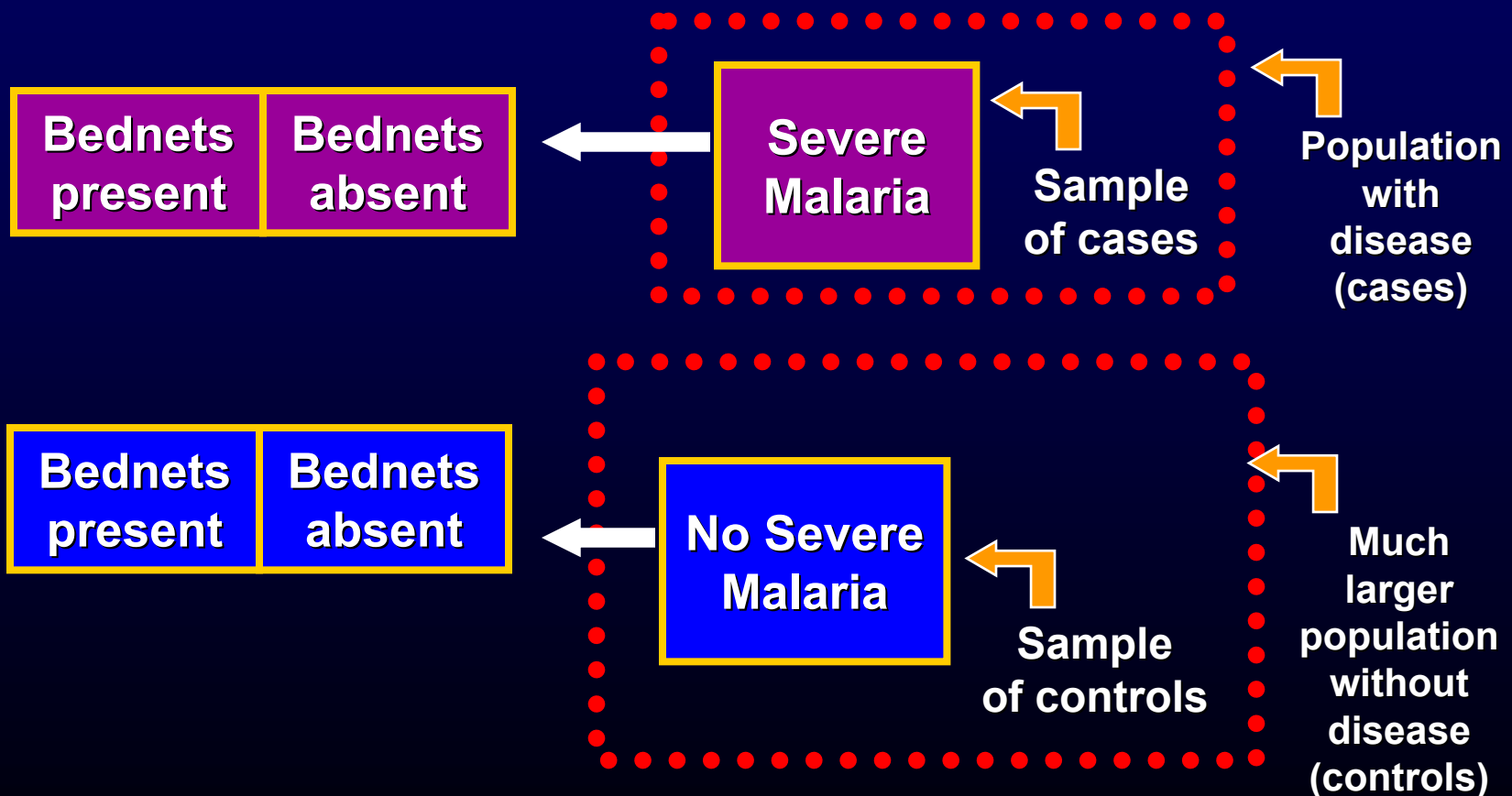


Start with a + b and c + d; then determine EXPOSURE

# Example of a Case-Control Design: Use Bednets and Severe Malaria in Children

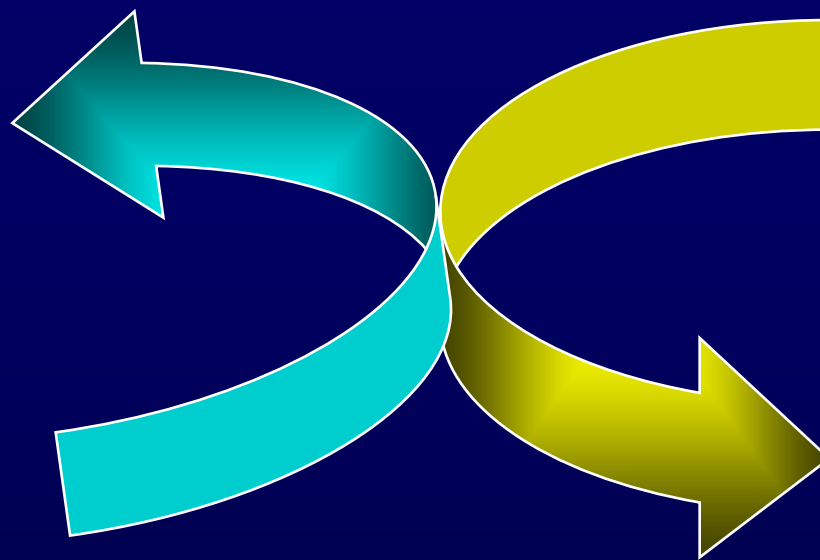
The Past or Present

The Present



**COHORT**  
.....▶

**TROHOC**  
◀.....



# Case-Control Study

## Some Advantages:

- ❖ Rare diseases
- ❖ Diseases with long latency period
- ❖ Fewer subjects
- ❖ Less Expensive
- ❖ Quicker to complete

# Case-Control Study

## Some Weaknesses:

- ❖ Recall bias
- ❖ Difficulty in selecting an appropriate control group
- ❖ Does not yield incidence rates in exposed and unexposed groups

# Case-Control Study

- ❖ Incidence rates cannot be calculated
- ❖ Hence, a relative risk cannot be calculated

**So what do we do?**

- ❖ Approximate the RR with an odds ratio

# RR Estimation in a Case-Control Study

		In Population Exposure		
		Yes	No	
Yes (Case)	<b>A</b>	<b>B</b>	<b>A + B</b>	
Outcome NO (Control)	<b>C</b>	<b>D</b>	<b>C + D</b>	
	<b>A + C</b>	<b>B + D</b>		

$$RR = \frac{A/(A + C)}{B/(B + D)}$$

If the rates are low ( $\approx < 5\%$ )

Then  $A/C \approx A/(A + C)$

$B/D \approx B/(B + D)$

$$RR \approx \frac{A/C}{B/D} = \frac{AD}{BC} = \text{odds ratio}$$

# RR Estimation in a Case-Control Study

	E	$\bar{E}$	
0 Cases	A 100	B 30	130
$\bar{0}$ Controls	C 2000	D 2000	4000
	2100	2030	

$$RR = 4.8\% / 1.5\% = 3.2$$

$$OR = 100 \cdot 2000 / 30 \cdot 2000 = 3.3$$

Cases	50	15	65
Controls	100	100	200

$$S_1 = .5$$

$$S_0 = .05$$

$$OR = 50 \cdot 100 / 15 \cdot 100 = 3.3$$

Incidence rates?

RR?

# Why Is A Case-Control Study Efficient?

## Cohort Study

	<b>E</b>	$\bar{E}$	
<b>Disease</b>	<b>15</b>	<b>10</b>	<b>25</b>
<b>No Disease</b>	<b>1000</b>	<b>2000</b>	<b>3000</b>
	<b>1015</b>	<b>2010</b>	<b><u>3025</u></b>

## Analogous Case Control Study

	<b>E</b>	$\bar{E}$	
<b>Cases (Disease)</b>	<b>15</b>	<b>10</b>	<b>25</b>
<b>Controls (No Disease)</b>	<b>10</b>	<b>20</b>	<b>30</b>
	<b>25</b>	<b>30</b>	<b><u>55</u></b>

# Why Is A Case-Control Study Efficient?

## Cohort Study

Study Size = 3025

RR = 3.0    95% CI (1.4 - 6.3)    p = .005

## Analogous Case-Control Study

Study Size = 55

OR = 3.0    95% CI (1.0 - 9.0)    p = .05

# Cohort Study

	Exposed	Unexposed	
Outcome	10	10	
No Outcome	4,990	19,990	
	5,000	20,000	Total Study Size 25,000

RR = 4.0

95% CI [1.7 – 9.6],

p = .001

# Case-Control Study

	Exposed	Unexposed	
Cases	10	10	20
Controls	10	40	50
			Total Study Size 70

OR = 4.0

95% CI [ 1.2 – 14.3],

p = .01

# Case-Control Examples

	<b>E</b>	<b><math>\bar{E}</math></b>	
<b>Cases</b>	15	10	<b>25</b>
<b>Controls</b>	10	20	<b>30</b>

<b># Controls</b>	<b>OR</b>	<b>95% C.I.</b>
30	3.0	.87 – 10.6
60	3.0	1.04 – 8.8
90	3.0	1.10 – 8.3
120	3.0	1.14 – 8.0

# Study Population

	$E$	$\bar{E}$
$0$	100	100,000
$\bar{0}$	200	1,000,000
	300	1,100,000

- Is a case-control study always efficient?
- Can a cohort study be more efficient?

# Case-Control

	Exposed	Unexposed		
0	10	10,000	10,100	Low Incidence of Exposure
$\bar{0}$	20	100,000	100,200	

# Cohort Study

	Exposed	Unexposed	
0	10	10	High Outcome Incidence
$\bar{0}$	20	100	
	30	110	

# Greater Chance of Bias in Case-Control Studies

- ❖ All of the relevant events, disease and exposure, have already occurred when the study begins
- ❖ Two potent sources of bias
  - ◆ Bias in ascertaining exposure
  - ◆ Selection of a control group

# Bias in Case-Control Studies

## Information Bias

- ❖ Information is gathered differently from cases and controls
- ❖ Difference related to risk factor
- ❖ Recall bias is most common

## Selection Bias

- ❖ Cases and controls are selected differently
- ❖ Difference in selection is related to risk factor

# Ascertainment Bias in Case-Control Studies

- ❖ Data collectors and participants can be prejudiced by knowledge of outcome
  - ◆ Especially if associations are alleged
- ❖ Try to ascertain exposure in an unbiased manner
  - ◆ Blinding
  - ◆ Visual aids to stimulate memory

# Case-Control Study

## Case Definition:

- ❖ Define objective criteria for reliable diagnosis of disease
- ❖ Know what population you are selecting cases from
  - ◆ Clinic, hospital, etc.
  - ◆ Population based

# Case-Control Study (Cont.)

## Control Selection:

- ❖ Controls should represent the population from which cases were selected
- ❖ Free of disease being studied
- ❖ Selection must be independent of exposure being studied
- ❖ Matching is not what it seems! Generally it is better not to match

# Case-Control Study of Risk Factors for AIDS

## Case Selection:

- ❖ MSM AIDS cases diagnosed in San Francisco during 1983-1984

## 2 Control Groups:

- ❖ MSM STD clinic patients
- ❖ MSM from same neighborhood as cases

# AIDS by Number of Partners

	> 100 Partners Per Year	0 – 5 Partners Per Year
AIDS Cases		
Controls		

**STD Clinic Controls (HIV-)**

**OR = 2.9**

**Neighborhood Controls (HIV-)**

**OR = 52.0**

Reference: Moss et. al. Am J Epidemiol 1987;125: 1035-47.

# STD Patients

	100+	0 - 5
AIDS	90	10
Controls	75	25

$$\text{OR} = \frac{90 \times 25}{10 \times 75} = 3.0$$

# Neighborhood

	100 +	0 - 5
AIDS	90	10
Controls	15	85

$$\text{OR} = \frac{90 \times 85}{10 \times 15} = 51.0$$

# NSAIDS → COLORECTAL CANCER

## Cases – Colorectal Cancer Patients

Hospital Controls	Exposure to NSAIDS	Impact on Odds Ratio
Arthritis	↑	Reduce
Peptic Ulcers	↓	Increase

# Advantages of Case-Control Studies

- ❖ **Case-control studies are useful and efficient for studying low frequency outcomes (i.e. approx. <5%)**
- ❖ **Case-control studies are useful for studying health problems with a long latent interval**
- ❖ **With rare outcomes, case-control studies are less time consuming and less expensive than cohort studies**

# **Disadvantages of Case-Control Studies Easier to Do . . . . .Wrong**

- ❖ **Prone to more selection bias than cohort studies: control group selection difficult**
- ❖ **Prone to more ascertainment bias than cohort: incomplete records or recall bias**
- ❖ **Cannot determine incidence rates**
- ❖ **By definition, pertinent to one outcome**
- ❖ **Cannot observe temporality of association**

# At a Minimum, for a Case-Control Study, in the Protocol Include

- ❖ Definition of cases
- ❖ Selection of the control group(s) [implications for selection bias]
- ❖ Selection and measurement of potential confounding factors
- ❖ Analogous endpoint: The ascertainment of exposure and the related procedures to minimize ascertainment bias (including whether any blinding would be attempted)

# **Randomized Controlled Trials**

---

**The methodologic standard  
of excellence for scientific  
experiments**

# Epidemiologic Study Designs

## ❖ Cohort

Exposure  Outcome

## ❖ Case-Control

Outcome  Exposure

## ❖ Randomized Controlled (Clinical) Trials (RCT)

# RCT PARADIGM

Population of Interest

Child  $\leq 5$  year presenting at hospital with severe malaria

Randomize

PTX

Placebo

Outcome Assessment  
Death within 7 days

**COHORT**



**TROHOC**



**RANDOMIZED**



**DEZIMODNAR**



# Assures Comparability

- ❖ **In observational studies, statistical methods allow investigators to control for confounding factors**
  - ◆ **Must be measured (ABLE)**

# Assures Comparability (cont.)

- ❖ No statistical method can achieve comparability on unknown or unmeasured factors in analysis phase
- ❖ Random allocation is the only known method to assure comparability

# Comparison of Results from Cohort Study and RCT

**Intervention: Exercise in men after M.I.**

**Outcome: Recurrent M.I.**

**Cohort: RR = .38    95% CI (0.18 - 0.80)    p = .006**

**RCT:    RR = 1.3    95% CI (0.73 - 2.2)    p = .20**

# **Hormone Replacement and Therapy Coronary Heart Disease (CHD)**

---

- ❖ Hormones decrease risk of CHD by 35% to 50%, according to 3 different meta-analysis of numerous observational studies**
- ❖ Especially strong for secondary prevention in women with CHD**

# Experimentation Trumps Observation

- ❖ RCT of hormone-therapy for secondary presentation of CHD
- ❖ Relative hazard = 0.99; 95% CI 0.80 - 1.22
- ❖ No effect of hormone therapy
- ❖ Recent Women's Health Initiative RCT in healthy women (JAMA 2002; 288: 321-333)
  - ◆ CHD: HR=1.29; 95% CI 1.02-1.63)

# Chemotherapy for Carcinoma of the Esophagus\*

- ❖ A meta-analysis of 8 non-randomized studies found a **68%** reduction in death (OR=0.32, 95% CI 0.24 - 0.42)
- ❖ A meta-analysis of 12 RCTs found a **4%** reduction in death  
(OR=0.96, 95% CI 0.75 - 1.22)

\*Bhansali et al. Ann Oncol 1996;7:355-9.

# **Trials of the Treatment of Acute M.I.**

	<b>RCTs <u>(57)</u></b>	<b>Non-Random <u>(43)</u></b>
<b>Difference in case-fatality rates at <math>p &lt; 0.05</math></b>	<b>8.8%</b>	<b>58%</b>
<b>Results favored treatment over controls (<math>p &lt; 0.05</math>)</b>	<b>60%</b>	<b>93%</b>
<b>Mean differences in the case-fatality rates</b>	<b>0.3%</b>	<b>10.5%</b>
	<b><u><math>\pm 0.8\%</math></u> <b>(<math>p &gt; 0.05</math>)</b></b>	<b><u><math>\pm 1.7\%</math></u> <b>(<math>p &lt; 0.001</math>)</b></b>

# **Systematic Review of Randomized vs Non-Randomized Evidence\***

- ❖ On average, non-randomized studies result in overestimates of effect**
- ❖ That bias can, however, go in either direction**
- ❖ That bias can be as large or larger than the effects of worthwhile interventions**

**\*Kunz R. and Oxman AD. BMJ 1998;317:1185-90.**

## EFFECT OF THE ANGIOTENSIN-CONVERTING-ENZYME INHIBITOR BENAZEPRIL ON THE PROGRESSION OF CHRONIC RENAL INSUFFICIENCY

**Abstract Background.** Drugs that inhibit angiotensin-converting enzyme slow the progression of renal insufficiency in patients with diabetic nephropathy. Whether these drugs have a similar action in patients with other renal diseases is not known. We conducted a study to determine the effect of the angiotensin-converting-enzyme inhibitor benazepril on the progression of renal insufficiency in patients with various underlying renal diseases.

The patients in each group were then randomly assigned to receive 10 mg of benazepril or placebo once daily. Randomization was balanced for disease severity at each center.

**New England Journal of Medicine**

# THE LANCET

## Double-blind trial of aspirin in primary prevention of myocardial infarction in patients with stable chronic angina pectoris

Clinical trials have demonstrated a prophylactic role for aspirin in myocardial infarction and in unstable angina pectoris. The Swedish Angina Pectoris Aspirin Trial (SAPAT) is a first prospective study of aspirin in stable angina pectoris.

**"After showing good tolerance of sotalol for at least three weeks the patients were randomised double blind to aspirin 75 mg daily (n=1009) or placebo (n=1026)."**

**Lancet**

# TREATMENT OF HIV INFECTION WITH SAQUINAVIR, ZIDOVUDINE, AND ZALCITABINE

**Abstract** *Background.* In patients with human immunodeficiency virus (HIV) infection, combined treatment with several agents may increase the effectiveness of antiviral therapy. We studied the safety and efficacy of saquinavir, an HIV-protease inhibitor, given with one or two nucleoside antiretroviral agents as compared with the safety and efficacy of a combination of two nucleosides alone.

**"The study (AIDS Clinical Trials Group protocol 229) was a randomized, double-blind, phase 2 trial of three treatment regimens"**

**New England Journal of Medicine**

## EFFECT OF THE ANGIOTENSIN-CONVERTING-ENZYME INHIBITOR BENAZEPRIL ON THE PROGRESSION OF CHRONIC RENAL INSUFFICIENCY

*Abstract Background.* Drugs that inhibit angiotensin-converting enzyme slow the progression of renal insufficiency in patients with diabetic nephropathy. Whether these drugs have a similar action in patients with other renal diseases is not known. We conducted a study to determine the effect of the angiotensin-converting-enzyme inhibitor benazepril on the progression of renal insufficiency in patients with various underlying renal diseases.

*Methods.* In a th... trial involving 583 patients with renal insufficiency caused by various... 1983 received pla...

**We conducted a prospective, double-blind, randomized study involving 49 European hospitals.**

**New England Journal of Medicine**

# Biocompatible membranes in acute renal failure: **prospective case-controlled study.**

Schiffl H, Lang SM, Konig A, Strasser T, Haider MC, Held E.

“The mortality of critically ill patients with acute renal failure has been halved through intervention by haemodialysis. However, several reports suggest that the course of the disorder may be prolonged by this procedure. Our **prospective randomised study** was done . . . .”

Lancet

# Randomized Trials Require Methodological Rigor

- ❖ Improperly conducted RCTs yield biased results
- ❖ Researchers must devote assiduous attention to design and conduct of RCTs
- ❖ Only properly conducted RCTs will fulfill their promise of minimizing bias
- ❖ Separate presentation later in the week

# Advantages of Randomized Trials

---

- ❖ **First and foremost, the only effective method known to control selection bias**
- ❖ **Controls confounding bias**
- ❖ **Facilitates effective blinding in some trials**
- ❖ **Theoretically attractive -many statistical methods assume random assignment**
- ❖ **Maintains advantages of cohort studies**

# Disadvantages of Randomized Trials

- ❖ **May be complex and expensive**
- ❖ **Difficult and expensive with low incidence outcomes**
- ❖ **May lack representativeness - volunteers may differ from population of interest**
- ❖ **Ethical challenges of experimental research**
- ❖ **Sometimes impossible or impractical**

**END**

---

**Thank You**

**The Fundamentals of International Clinical Research**  
**Workshop**  
**Dar es Salaam, Tanzania**  
**January 2010**

**Case Scenarios**

- **Case 1:** Dr. Gates would like to evaluate whether the presence of helminth infection at the time of vaccination with conjugate pneumococcal vaccine impairs antibody responses in children ages 4-8 years in Cameroon. It is postulated that the children with helminth infections will have a Th2 shift in cytokines resulting in impaired antibody responses to conjugate vaccines.

Previous studies of children in the districts of interest show a 40% prevalence of helminth infections. Participants in prior studies have displayed excellent retention yielding over 95% follow up rates. Dr. Gates estimates that approximately 90% of children in Cameroon without helminth infections would develop adequate antibody titers. She worries that only 70% of children with helminth infections would develop adequate antibody titers.

- **Case 2:** Dr. Michigan working in Malawi would like to determine if adjuvant treatment with 2' Hydroxychalcone in children with cerebral malaria reduces the incidence of death within 14 days of admission to hospital. Chalcones reduce expression of endothelial adhesion molecules which are up regulated in children with severe malaria. In vitro data suggest that 2'-hydroxychalcone decreases transcription of ICAM-1, VCAM-1 and E-selectin genes.

Since all the patients would be hospitalized, the study should realistically yield 100% follow up. Approximately 30% of the children admitted to hospital with cerebral malaria die within 14 days. Dr. Michigan postulates that 2'Hydroxychalcone will lead to a 50% reduction in the incidence of death.

- **Case 3:** Drs. Fixit and Royalty would like to determine if positive hepatitis serology, consumption of alcohol, and/or acetaminophen (all with prevalences of 20 to 50% in the population of interest) contribute to the development of portal hypertension in individuals infected with *S. japonicum* in China. They believe that a long latency interval exists between all the examined exposures and the putative association with the development of portal hypertension.

They estimate that the incidence of portal HTN in those with *S. japonicum* is approximately 1-2% in the population of interest. Local health departments have excellent records on the population of interest dating back for 15 years.

- **Case 4:** Investigators in Egypt, Yemen, and Kenya have reviewed the Rift Valley Fever (RVF) literature and believe that antibody inversely correlates with disease outcome. They hypothesize that low antibody production contributes to higher mortality. Further, they hypothesize that IL-12 will increase RVF specific antibody and result in a decreased mortality from RVF compared to a group not receiving IL-12.

They estimate that mortality in those with RVF not receiving IL-12 is approximately 3-4%. In prior studies at their research sites of RVF, investigators have displayed an excellent ability to follow-up participants. They would anticipate being able to achieve about a 98% retention rate in an RVF study.

- **Case 5:** The Malaria Treatment Research Network has decided to evaluate the efficacy of Falgone, a new 8-amino quinolone, in children with moderately severe malaria.

The current standard drug choice in such situations has a treatment failure rate of about 30% at 28 days, based on a few studies in the published literature, with most failures being LTF (late treatment failure). Minimal side effects are associated with the current standard drug of choice.

Based on coverage, Falgone shows great promise. However, some of the untoward effects thought to be related to the drug in the phase II study include neutropenia, transient blurred vision, and proteinuria.