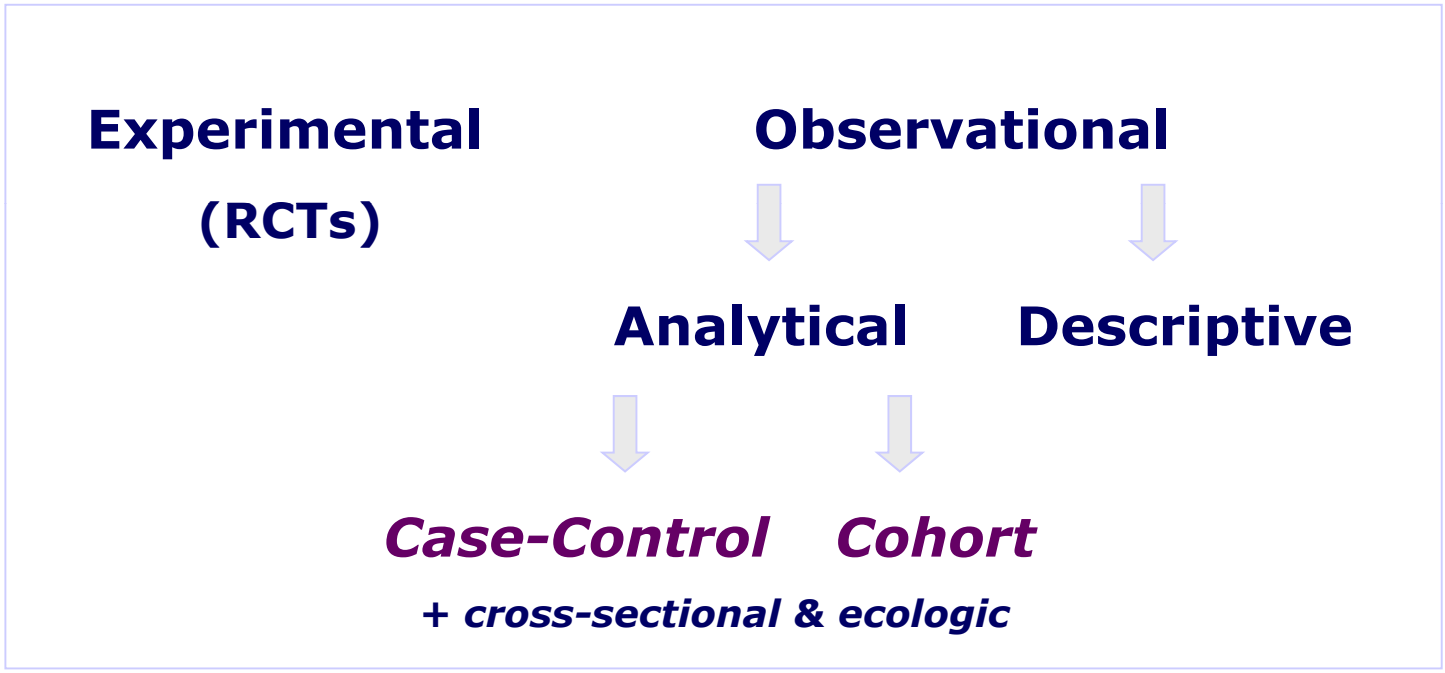
Epidemiologic Study Designs

Epidemiologic Study Designs



Epidemiologic Study Designs

Descriptive studies

Examine patterns of disease

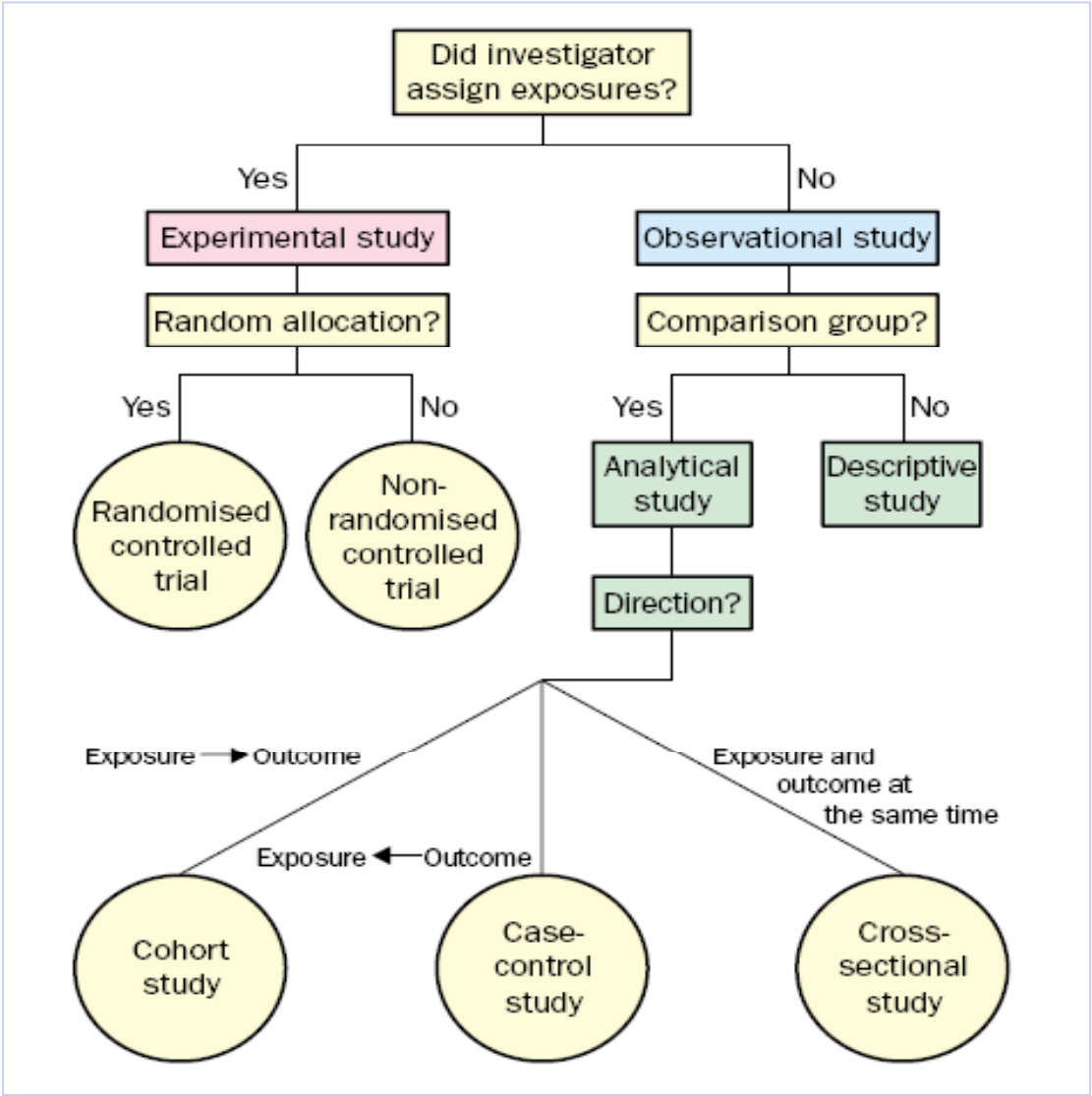
Analytical studies

Studies of suspected causes of diseases

Experimental studies

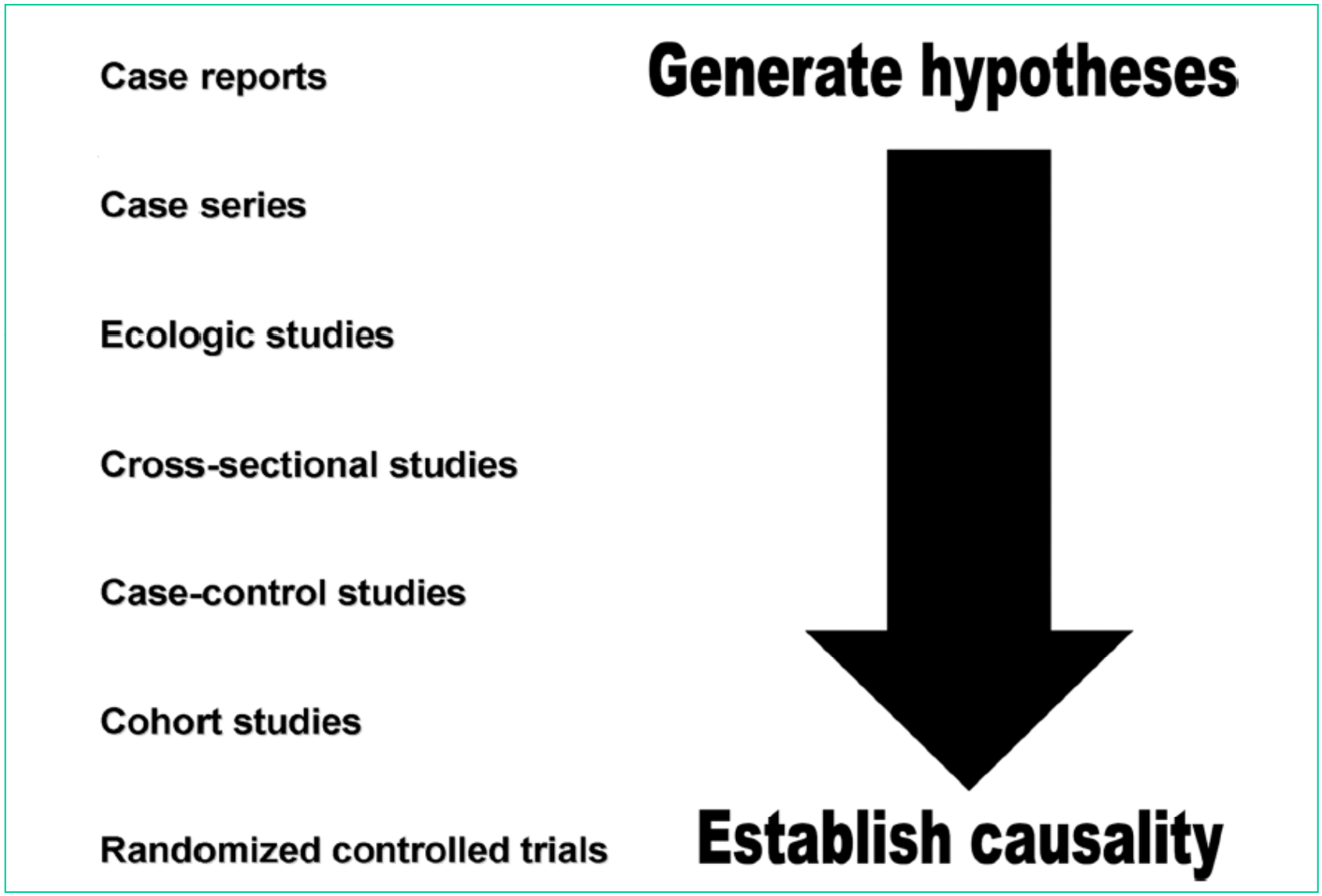
Compare treatment modalities

Epidemiologic Study Designs



Grimes & Schulz, 2002 (www)

Hierarchy of Epidemiologic Study Design



Tower & Spector, 2007 (www)

Observational Studies

(no control over the circumstances)

- Descriptive: Most basic demographic studies
- Analytical: Comparative studies testing an hypothesis
  - * cross-sectional

    (a snapshot; no idea on cause-and-effect relationship)
  - * cohort

    (prospective; cause-and-effect relationship can be inferred)
  - * case-control

    (retrospective; cause-and-effect relationship can be inferred)

Epidemiologic Study Designs

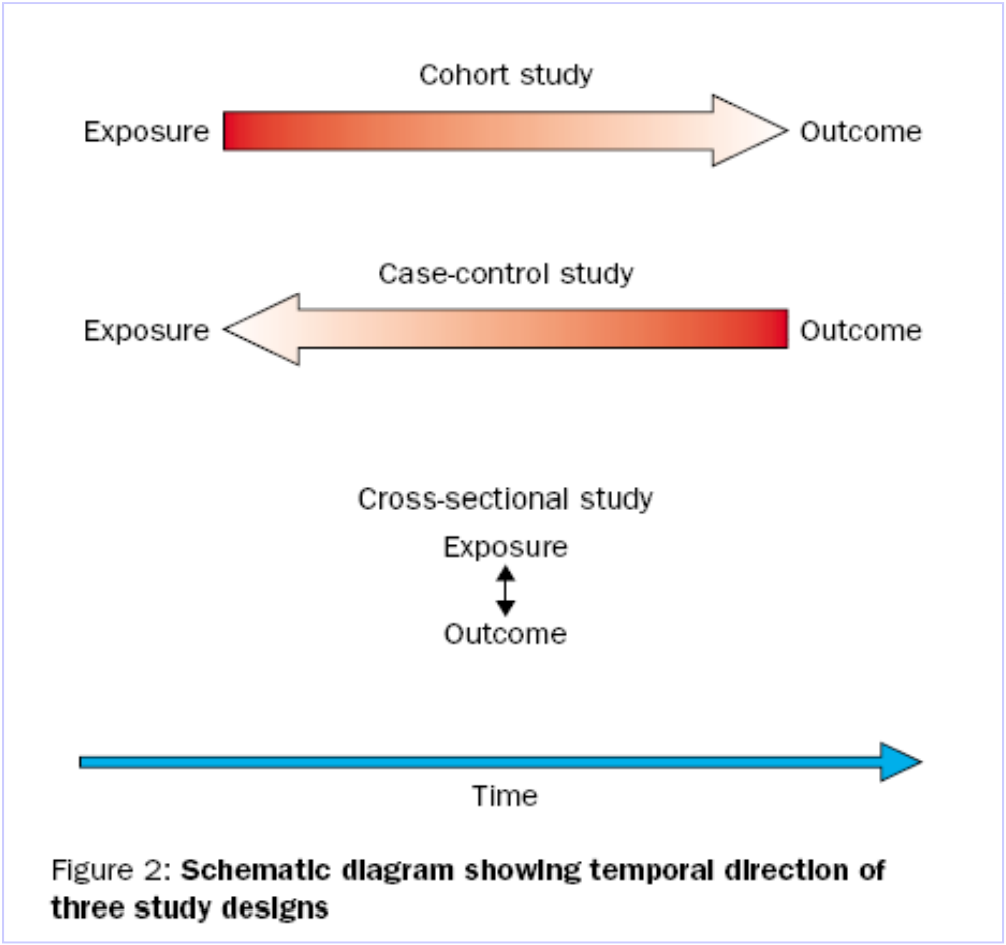


Figure 2: **Schematic diagram showing temporal direction of three study designs**

Grimes & Schulz, 2002 (www)

Analytical Studies

(comparative studies testing an hypothesis)

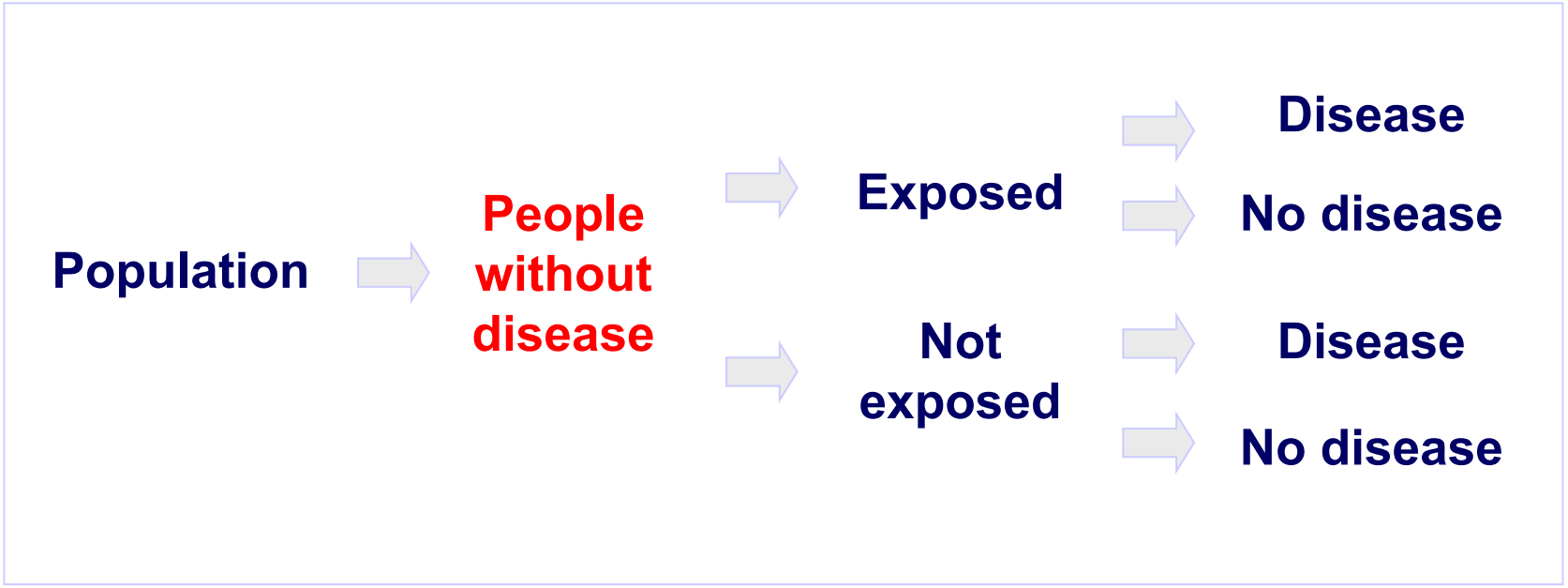
* **cohort** (prospective)

  Begins with an exposure (smokers and non-smokers)

* **case-control** (retrospective - trohoc)

  Begins with outcome (cancer cases and healthy controls)

Cohort Studies



Examples of Cohort Studies

* ***Framingham Heart Study*** *(www)*

* ***NHANES Studies*** *(www)*

* ***MACS*** *(www)*

* ***Physicians' Health Study*** *(www)*

* ***Nurses' Health Study*** *(www)*

* ***ALSPAC*** *(www)*

Advantages of Cohort Studies

- Can establish population-based incidence
- Accurate relative risk (risk ratio) estimation
- Can examine rare exposures (asbestos > lung cancer)
- Temporal relationship can be inferred (prospective design)
- Time-to-event analysis is possible
- Can be used where randomization is not possible
- Magnitude of a risk factor's effect can be quantified
- Selection and information biases are decreased
- Multiple outcomes can be studied
  (smoking > lung cancer, COPD, larynx cancer)

Disadvantages of Cohort Studies

- Lengthy and expensive
- May require very large samples
- Not suitable for rare diseases
- Not suitable for diseases with long-latency
- Unexpected environmental changes may influence the association
- Nonresponse, migration and loss-to-follow-up biases
- Sampling, ascertainment and observer biases are still possible

Presentation of cohort data: Population at risk

Does HIV infection increase risk of developing TB among a population of drug users?

	Population (follow up 2 years)	Cases
HIV +	215	8
HIV -	289	1

Source: Selwyn et al., New York, 1989

Does HIV infection increase risk of developing TB among drug users?

Exposure	Population (f/u 2 years)	Cases	Incidence (%)	Relative Risk
HIV +	215	8	3.7	11
HIV -	298	1	0.3	

Presentation of cohort data: Person-years at risk

Tobacco smoking and lung cancer, England & Wales, 1951

	Person-years	Cases
Smoke	102,600	133
Do not smoke	42,800	3

Source: Doll & Hill

Presentation of data: Various exposure levels

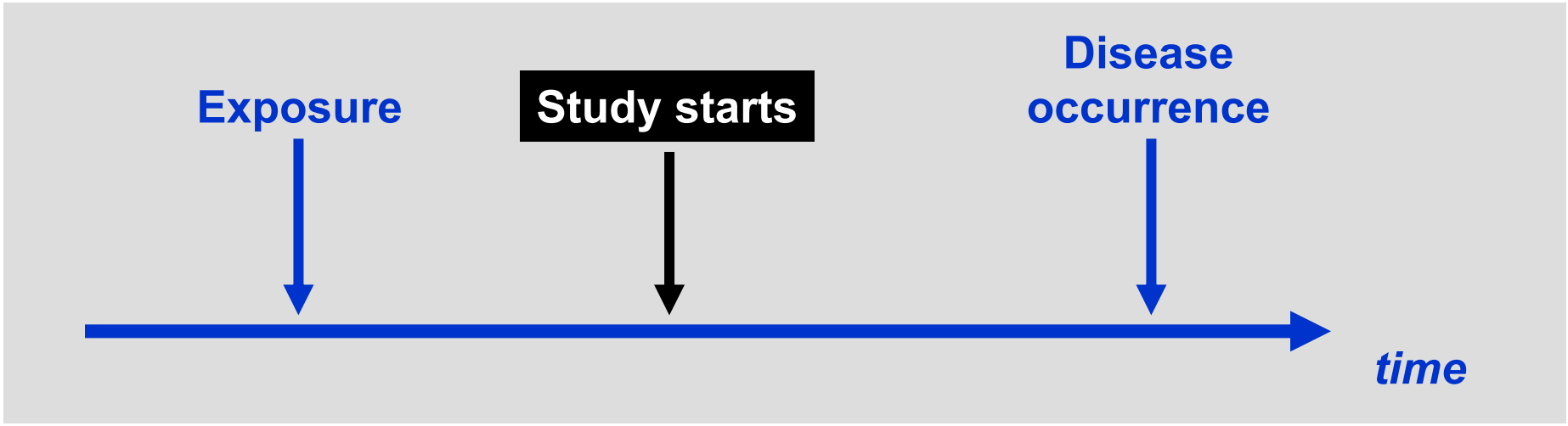
Daily number of cigarettes smoked	Person-years at risk	Lung cancer cases
> 25	25,100	57
15 - 24	38,900	54
1 - 14	38,600	22
none	42,800	3

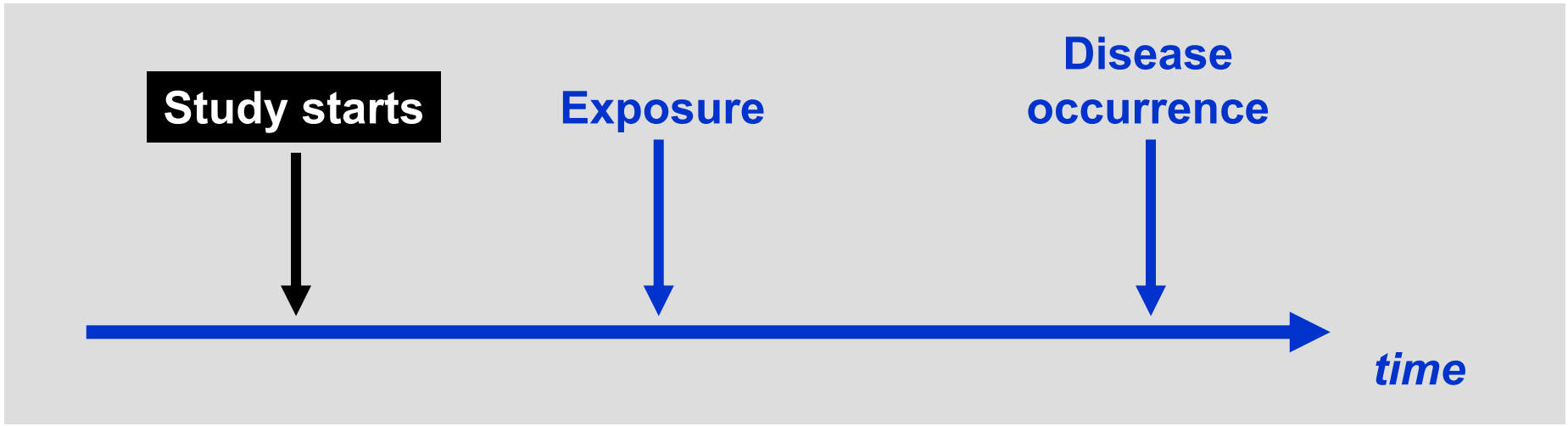
Cohort study: Tobacco smoking and lung cancer, England & Wales, 1951

Cigarettes smoked/d	Person-years at risk	Cases	Rate per 1000 p-y	Rate ratio
> 25	25,100	57	2.27	32.4
15 - 24	38,900	54	1.39	19.8
1 - 14	38,600	22	0.57	8.1
none	42,800	3	0.07	Ref.

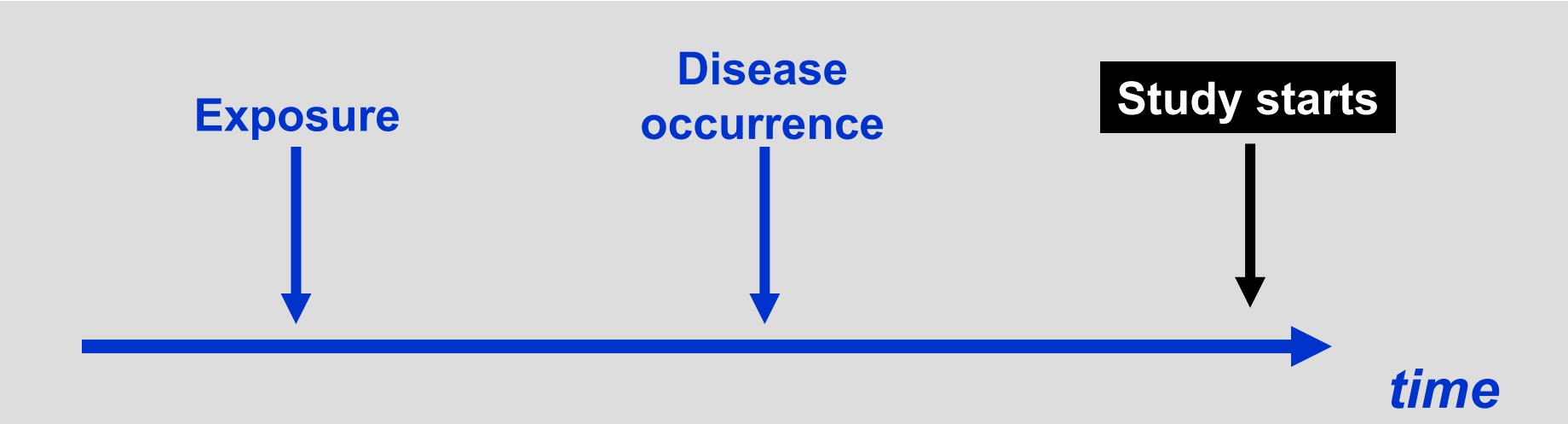
Source: Doll & Hill

Prospective cohort study





Retrospective cohort studies



Cohort Studies

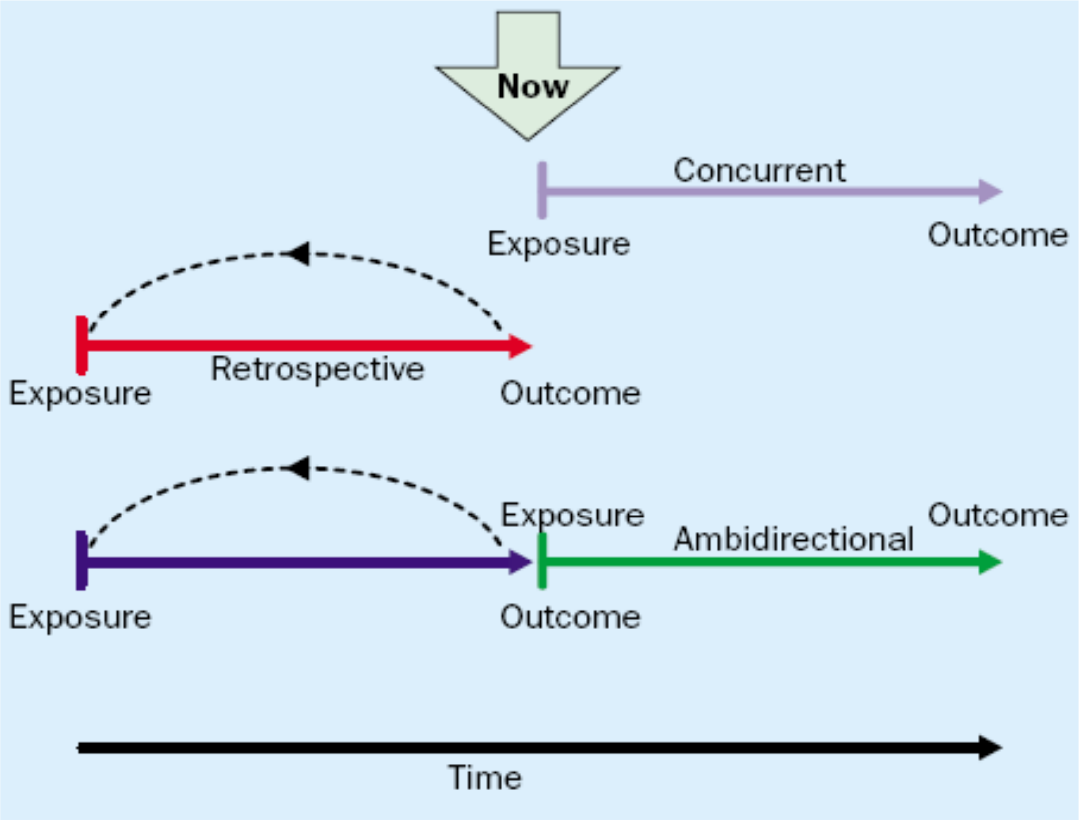


Figure 2: **Schematic diagram of concurrent, retrospective, and ambidirectional cohort studies**

Grimes & Schulz, 2002 (www) (PDF)

Cohort Studies

Panel 2: **Features to look for in a cohort study**

How much selection bias was present?

1 Were only people at risk of the outcome included?
1 Was the exposure clear, specific, and measurable?
1 Were the exposed and unexposed groups similar in all important respects except for the exposure?

What steps were taken to minimise information bias?

1 Was the outcome clear, specific, and measurable?
1 Was the outcome identified in the same way for both groups?
1 Was determination of outcome made by an observer blinded as to treatment?

How complete was the follow-up of both groups?

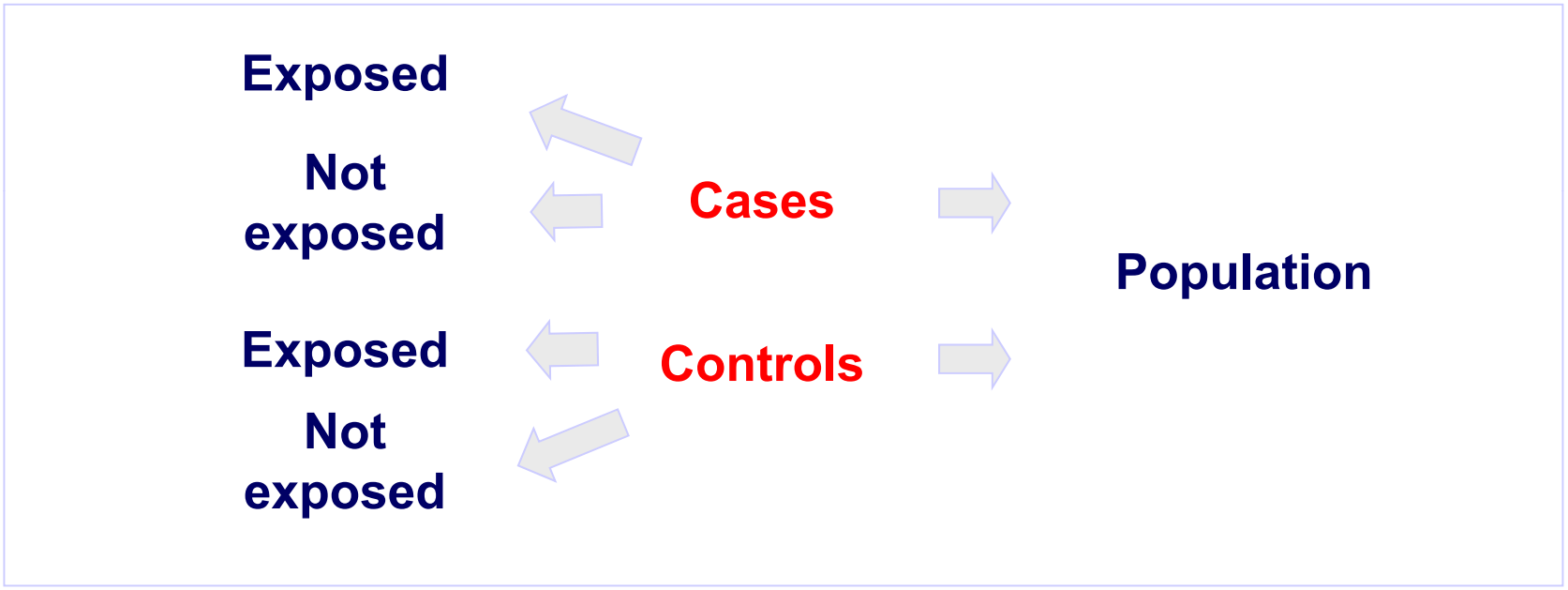
1 What efforts were made to limit loss to follow-up?
1 Was loss to follow-up similar in both groups?

Were potential confounding factors sought and controlled for in the analysis?

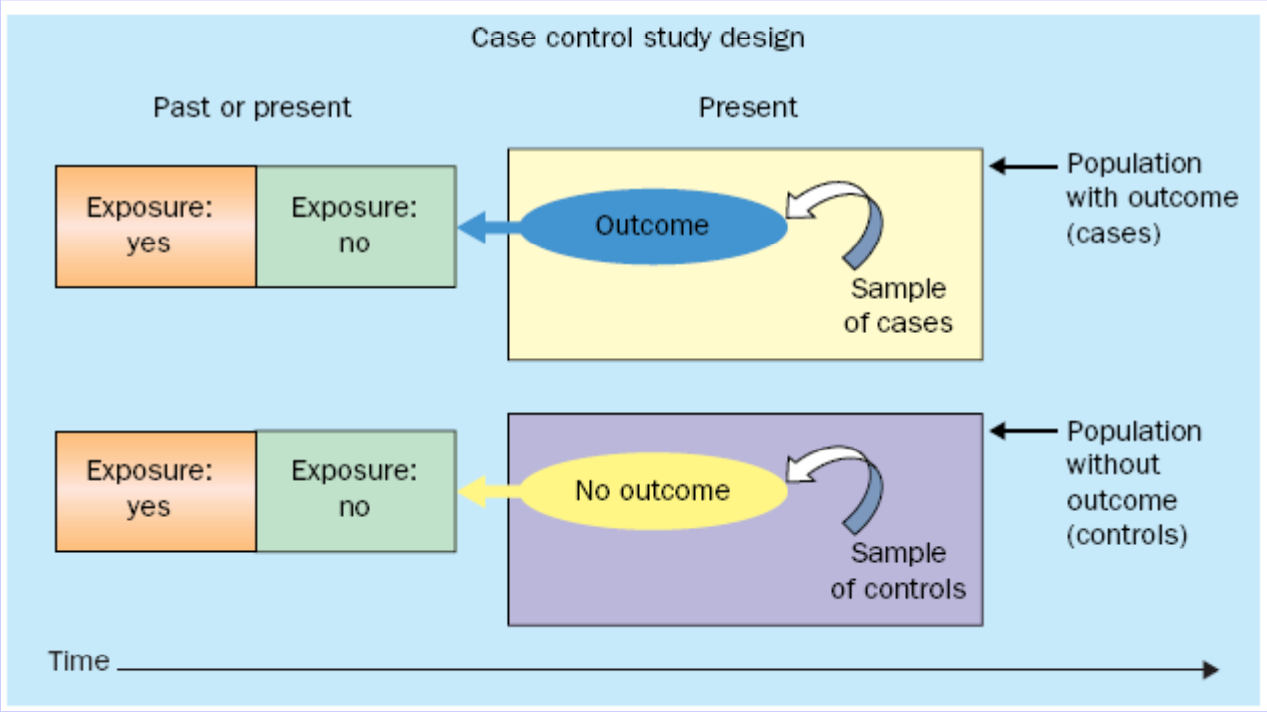
1 Did the investigators anticipate and gather information on potential confounding factors?
1 What method(s) were used to assess and control for confounding?

Grimes & Schulz, 2002 (www) (PDF)

Case-Control Studies



Case-Control Studies



Schulz & Grimes, 2002 (www) (PDF)

Advantages of Case-Control Studies

- Cheap, easy and quick studies
- Multiple exposures can be examined
- Rare diseases and diseases with long latency can be studied
- Suitable when randomization is unethical (alcohol and pregnancy outcome)

Disadvantages of Case-Control Studies

- Case and control selection troublesome
- Subject to bias (selection, recall, misclassification)
- Direct incidence estimation is not possible
- Temporal relationship is not clear
- Multiple outcomes cannot be studied
- If the incidence of exposure is high, it is difficult to show the difference between cases and controls
- Not easy to estimate attributable fraction
- Reverse causation is a problem in interpretation - especially in molecular epidemiology studies

Case-Control Studies: Potential Bias

Panel 2: **Introduction of bias through poor choice of controls**

Cases	Control selection	Non-representativeness	Selection bias
Colorectal cancer patients admitted to hospital	Patients admitted to hospital with arthritis	Controls probably have high degrees of exposure to NSAIDs	Would spuriously **reduce** the estimate of effect (odds ratio)
Colorectal cancer patients admitted to hospital	Patients admitted to hospital with peptic ulcers	Controls probably have low degrees of exposure to NSAIDs	Would spuriously **increase** the estimate of effect (odds ratio)

NSAIDs=non-steroidal anti-inflammatory drugs.

Schulz & Grimes, 2002 (www) (PDF)

Cause-and-Effect Relationship

Temporal sequence
Did exposure precede outcome?

Strength of association
How strong is the effect, measured as relative risk or odds ratio?

Consistency of association
Has effect been seen by others?

Biological gradient (dose-response relation)
Does increased exposure result in more of the outcome?

Specificity of association
Does exposure lead only to outcome?

Biological plausibility
Does the association make sense?

Coherence with existing knowledge
Is the association consistent with available evidence?

Experimental evidence
Has a randomised controlled trial been done?

Analogy
Is the association similar to others?

Grimes & Schulz, 2002 (www) (PDF)

Cause-and-Effect Relationship

Panel 1: **What to look for in observational studies**

Is selection bias present?
In a cohort study, are participants in the exposed and unexposed groups similar in all important respects except for the exposure?

In a case-control study, are cases and controls similar in all important respects except for the disease in question?

Is information bias present?
In a cohort study, is information about outcome obtained in the same way for those exposed and unexposed?

In a case-control study, is information about exposure gathered in the same way for cases and controls?

Is confounding present?
Could the results be accounted for by the presence of a factor—eg, age, smoking, sexual behaviour, diet—associated with both the exposure and the outcome but not directly involved in the causal pathway?

If the results cannot be explained by these three biases, could they be the result of chance?
What are the relative risk or odds ratio and 95% CI?[11,12]

Is the difference statistically significant, and, if not, did the study have adequate power to find a clinically important difference?[13,14]

If the results still cannot be explained away, then (and only then) might the findings be real and worthy of note.

Grimes & Schulz, 2002 (www) (PDF)

Epidemiologic Association / Impact Measures

(Absolute Risk) (AR)

Relative Risk (Risk Ratio) (RR)

Odds Ratio (OR)

Phi coefficient / Cramer's V / Contingency coefficient

Attributable Fraction (AF)

Attributable Risk (AR)

Relative Risk Reduction (RRR)

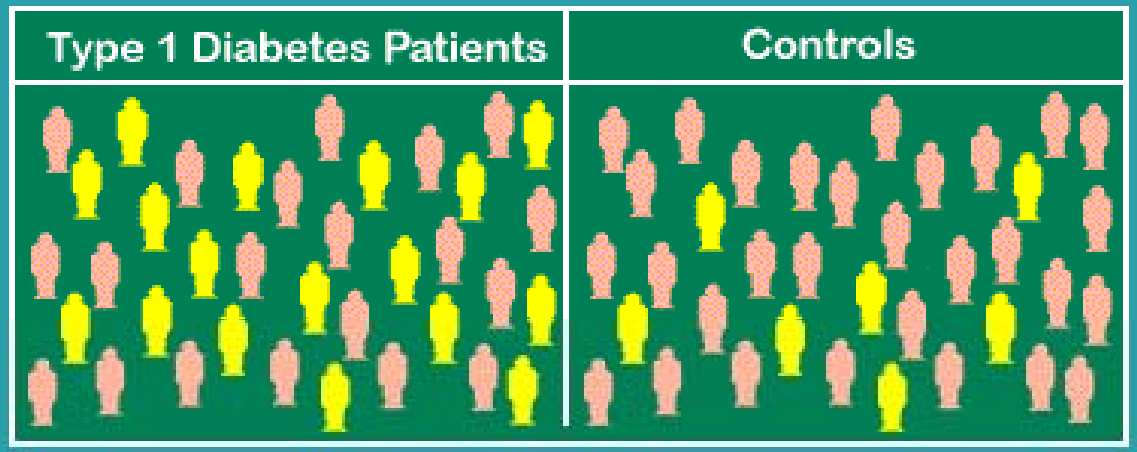
Absolute Risk Reduction (ARR)

Number Needed to Treat (NNT)

Measures of test accuracy:

Sensitivity, specificity, positive and negative predictive value (PPV, NPV)

Association Studies



Genotype	Type 1	Controls	Total
HLA DR4	17	7	24
NON-HLA DR4	20	30	50
	37	37	

$X^2_{.05} = 5.377$

$p < 0.025$

= HLA DR4

= non-HLA DR4

Odds Ratio: 3.6
95% CI = 1.3 to 10.4

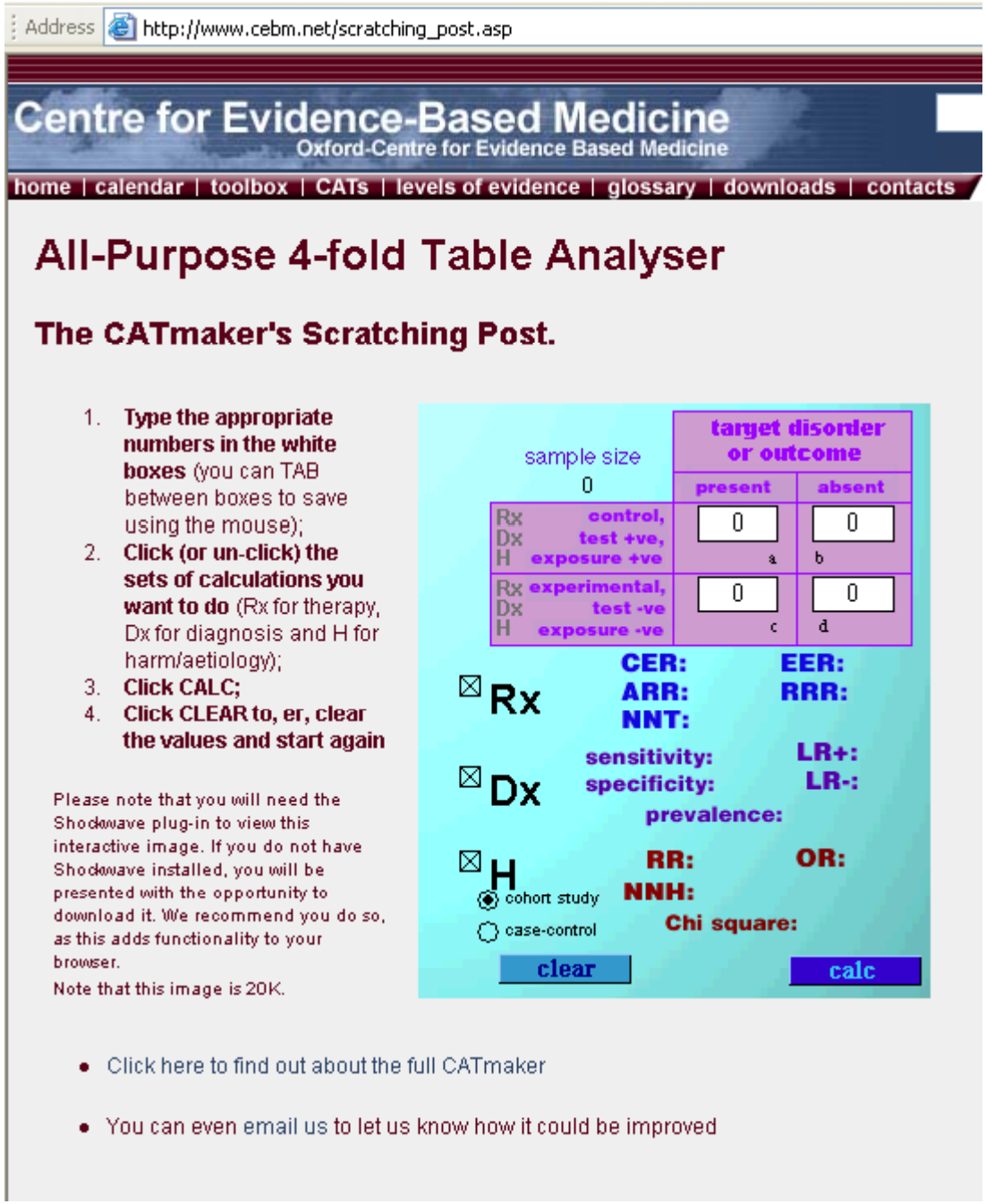
ROCHE Genetic Education (www)

Genotype	Type 1	Controls	Total
HLA DR4	17	7	24
NON-HLA DR4	20	30	50
	37	37	

a = 17
b = 20
c = 7
d = 30

OR = ad / bc = 17*30 / 20*7 = 3.6
RR = (a/(a+c)) / (b/(b+d)) = (17/24)/(20/50) = 1.8

EBM toolbox (www)
EpiMax Table Calculator (www)



EBM toolbox (www)

EpiMax Table Calculator

Epidemiology & Lab Statistics from Study Counts
With Chi Square, NNT & "Cost to Treat" Estimates

[For Demonstration Only-Not for Official Use]

Clinical & Economic Software Solutions	
Health Decision Strategies, LLC	Princeton, New Jersey USA

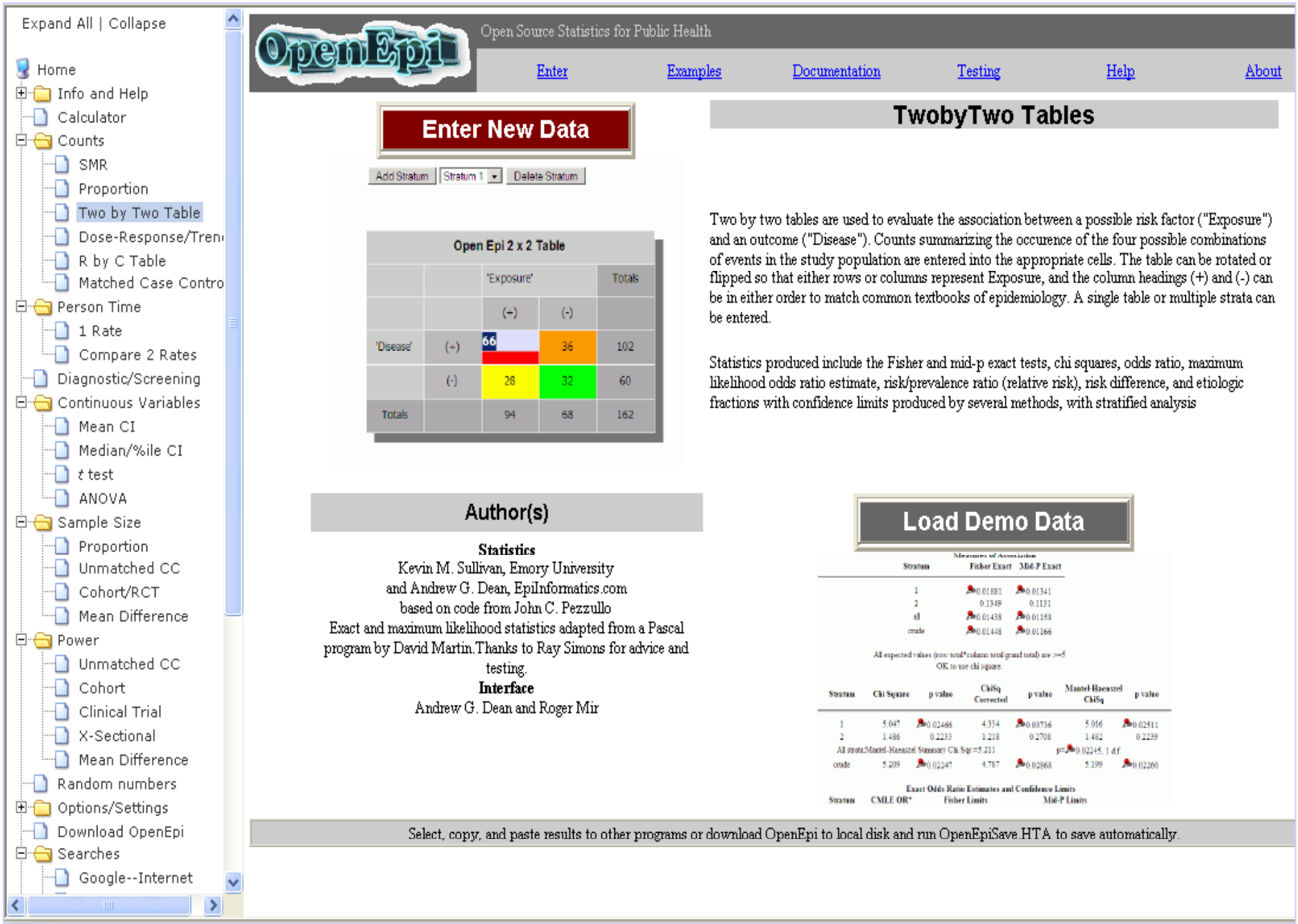
Data Entry Page

Instructions: Using 2x2 study data, you can change the "Title" and fill in the four center cells in the table below (the cells in blue) and if you wish, enter a "Cost Per Person" value. Hit the "Calculate" button to see the estimated results.

(Results generated will appear in the boxes outside and below the center cells.)

		Target Disorder or Outcome		
	Analysis Title:	Present	Absent	
	TPA vs Strepto: Fiction	Case	Control	
		True Positive(a)	False Positive(b)	a + b
Rx Dx H	◉ Control Group ○ Diag. Test positive ○ Exposed to Risk Factor	100	900	
		False Negative(c)	True Negative(d)	c + d
Rx Dx H	-- Experimental Group -- Diag. Test negative -- Not Exposed to Risk	90	910	
	Incremental Cost Per Person (CPP) Per Duration	a + c	b + d	a+b+c+d
	$ 2000			

EpiMax Table Calculator (www)



Open-Epi Calculator (www)

Epidemiologic Study Designs

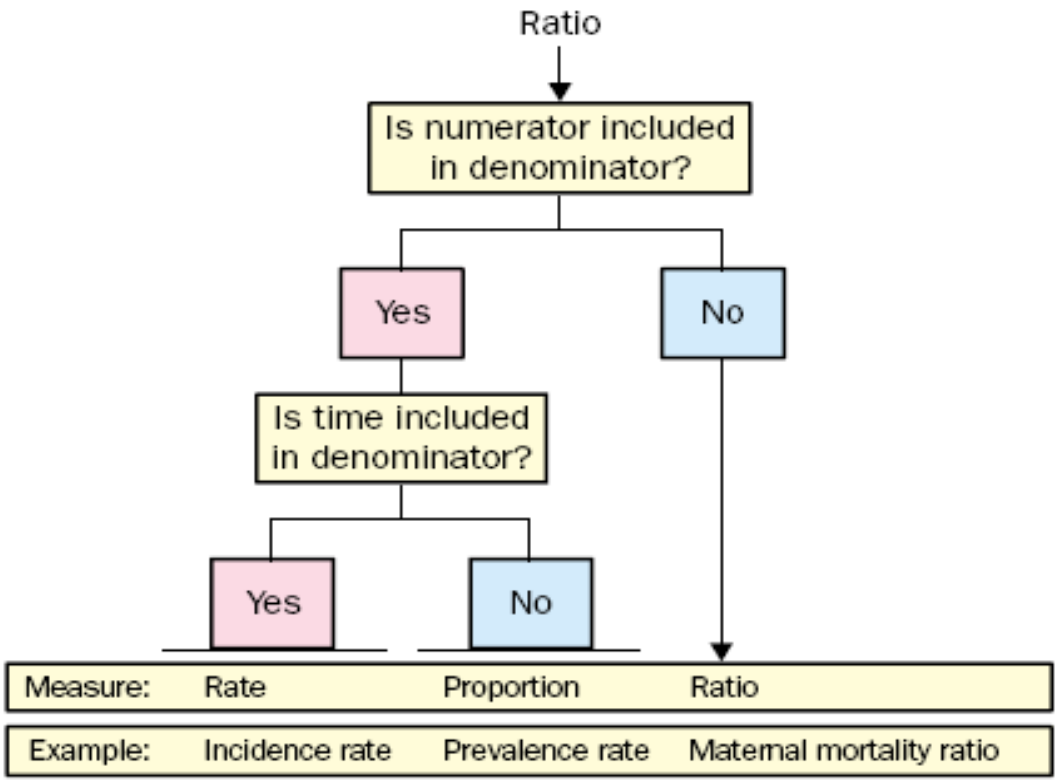


Figure 3: **Algorithm for distinguishing rates, proportions, and ratios**

Grimes & Schulz, 2002 (www)

Sources of Error in Epidemiologic Studies

Random error

Bias

Confounding

Effect Modification

Reverse Causation

Sources of Error in Epidemiologic Studies

Random error

Large* sample *size, replication

Bias

Be careful

Confounding

Effect Modification

Reverse Causation

Confounding can be controlled by:

- **Randomization**: assures equal distribution of confounders between study and control groups

- **Restriction**: subjects are restricted by the levels of a known confounder

- **Matching**: potential confounding factors are kept equal between the study groups

- **Stratification** for various levels of potential confounders

- **Multivariable analysis** (does not control for *effect modification*)

Effect modification can be assessed by:

- Stratification for various levels of potential confounders

- Multivariable analysis (by assessing interaction)

More importantly, *NOT* by adjustment in multivariable analysis

Reverse causation can be assessed by:

- Mendelian randomization