

# Chapter 5 Cohort study

**Lina Wang**



**Department of Epidemiology and Biostatistics  
School of Public Health  
Southeast University  
Phone: 025-83272569 (o)  
E-mail: [lnwang@seu.edu.cn](mailto:lnwang@seu.edu.cn)**

# Patient Profile

- ◆ The mother, a 28-year-old primigravida, had experienced elevated blood pressure during the pregnancy and the pregnancy had continued 2 weeks past the expected date of delivery.
- ◆ At the time of delivery, the male newborn was limp and cyanotic, and had no spontaneous respiratory effort and a heart rate of only **50** beats /minute. The baby did not grimace, cough, or sneeze.





## Apgar score for evaluation of neonatal asphyxia.

Sign	Score		
	0	1	2
Heart rate(beats/minute)	Absent	<100	>100
Respiration	Absent	Slow, irregular	Regular, crying
Muscle tone	Limp	Slow, flexion	Active motor
Color	Blue, pale	Body pink, extremities blue	Completely pink
Reflex response to catheter in nostril	None	Grimace	Cough, sneeze

The values for each of the five categories are added to yield a result from 0 to 10.



# Patient Profile

- ◆ The 5-minute Apgar score improved only to **2**.
- ◆ The 10-minute Apgar score remained depressed at **3**.
- ◆ With aggressive medical management in ICU, the 3100-g neonate continued to improve without evidence of **acute neurologic complications**.

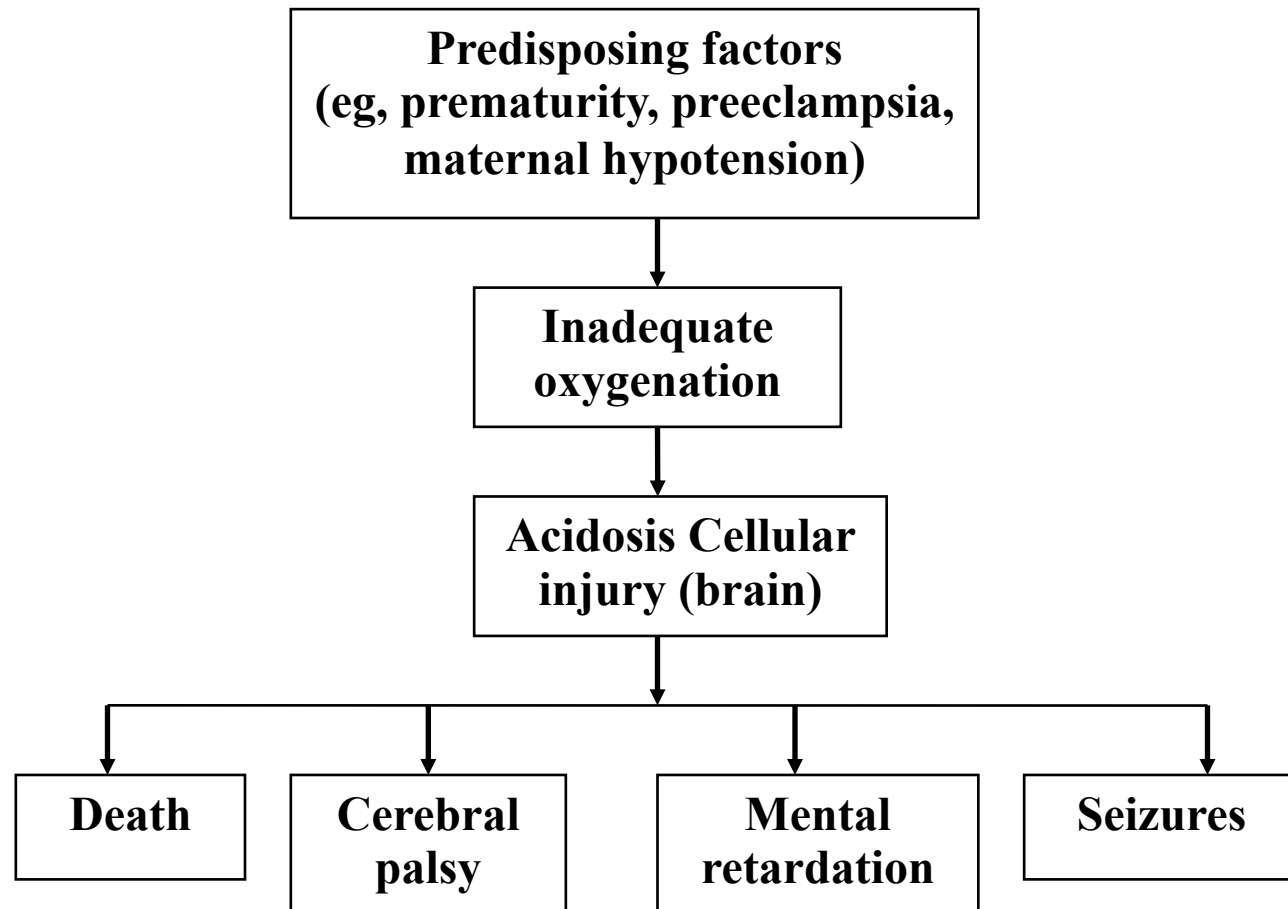


# Clinical background

- **Perinatal asphyxia can be defined as fetal hypoxia during labor and delivery.**
- **The causes of perinatal asphyxia are not completely understood, but a number of factors have been associated with hypoxia during labor and delivery, including preeclampsia or eclampsia, maternal hypotension, placental insufficiency, and prematurity.**



## Schematic representation of the pathogenesis of perinatal asphyxia.





# Study Design

## Questions from the Parents

- **What the future will bring?**
- **Will their son develop normal mental capacity?**
- **Will he have physical disabilities?**



# The outcomes of infants with severe perinatal asphyxia

- **Dismal outcomes, including mental retardation, physical disabilities and death.**
- **Normal neurologic development and excellent school performance.**

**What are the reliable evidences?**

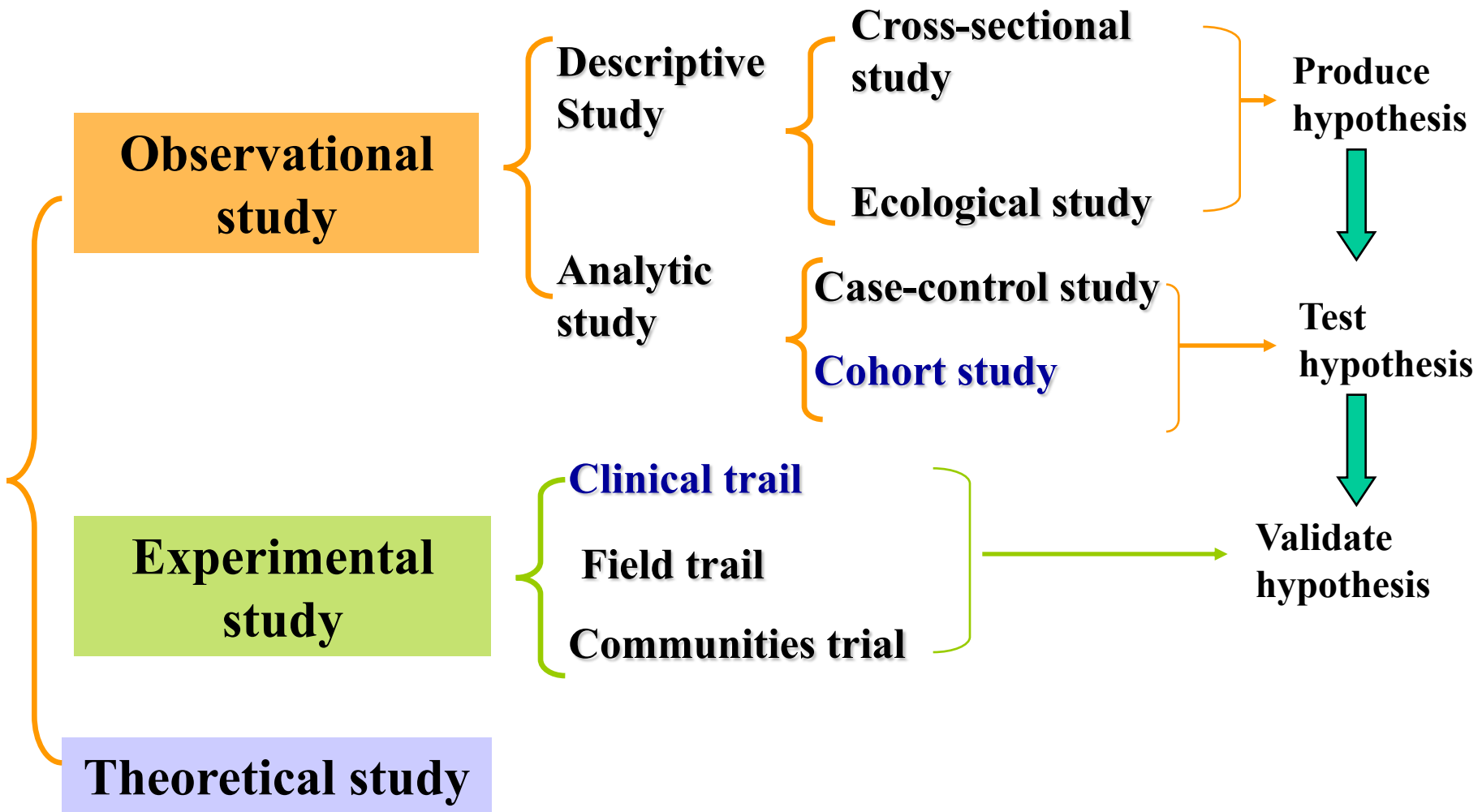


**We need to test the hypothesis:**

**the risk factors (lower Apagar scores) are associated with the increased risk of the dismal outcomes.**



# Epidemiology Study Methods





# Concept Map

**Descriptive  
Epidemiology**



**Hypothesis**



**Analytical  
Epidemiology**

**RCT**



**Judging  
Causality**



**Prevention  
Strategies**

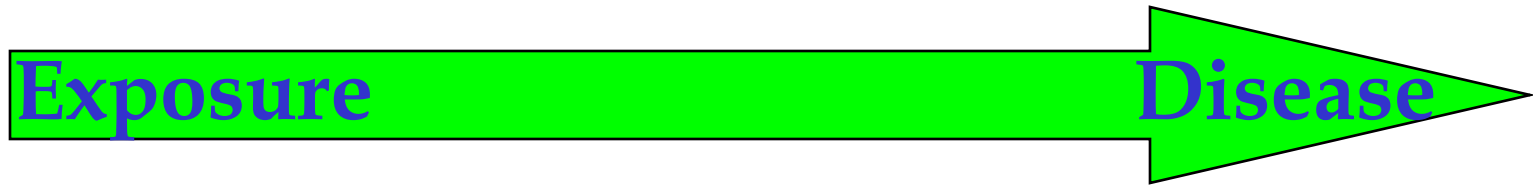


## Analytic Epidemiology

The aspect of epidemiology concerned with the search for health-related **causes and effects**. Uses **comparison groups**, which provide baseline data, to **quantify the association between exposures and outcomes**, and **test hypotheses about causal relationships**.



Time



*Cohort study*



*Case control study*



- A major limitation of **cross-sectional surveys** and **case-control studies** is difficulty in determining if exposure or risk factor preceded the disease or outcome.



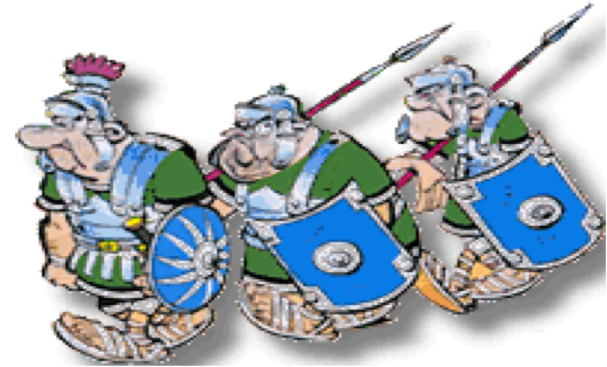
# Cohort Study



## Key Point:

- Presence or absence of risk factor (exposure) is determined **prior to** the observation of disease status.

## Cohort



- One of 10 divisions of a **Roman legion**
- Group of individuals
  - **sharing same experience**
  - **followed up for specified period of time**



## Cohort study

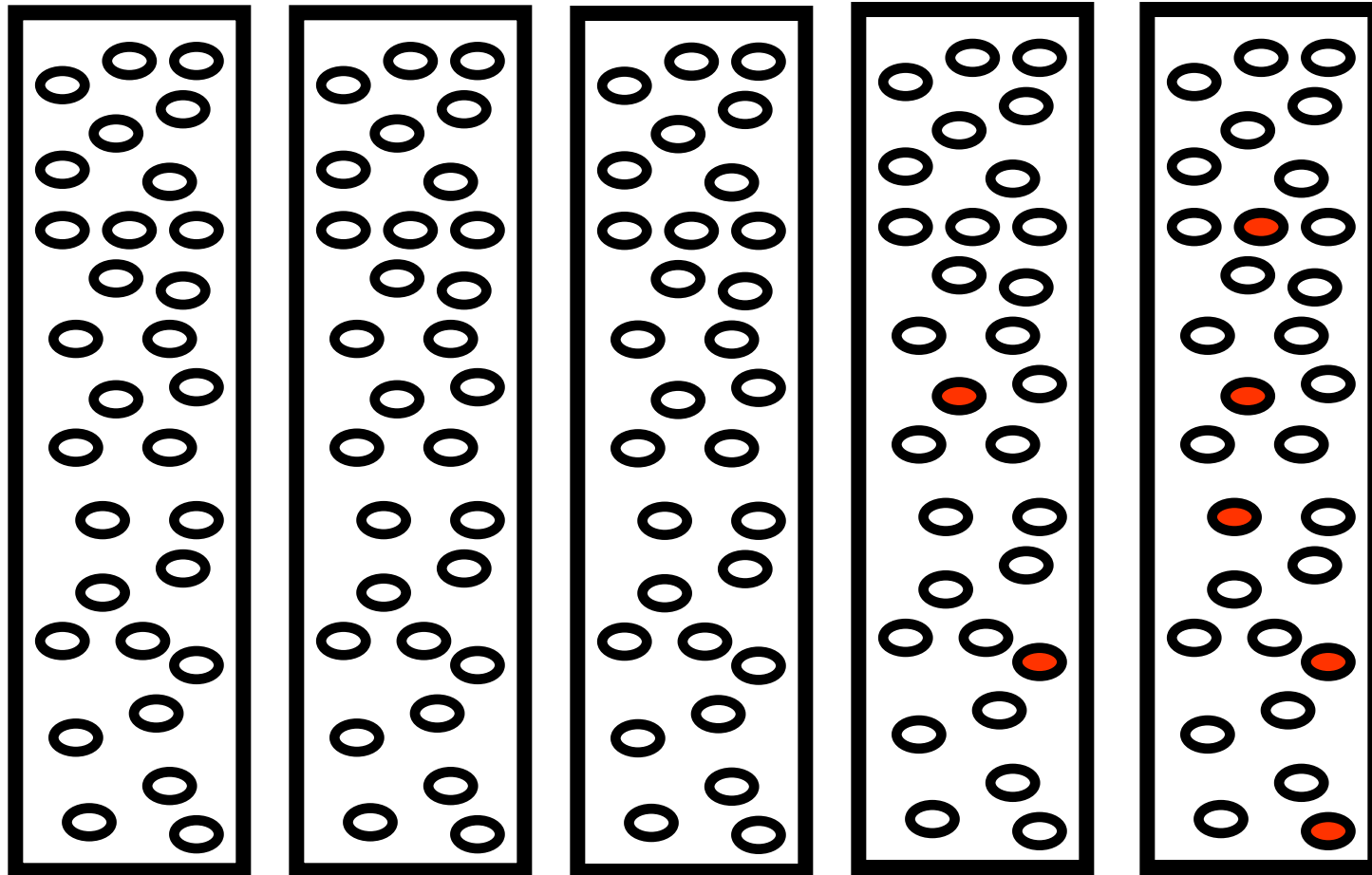
- A study in which a group or groups of individual (i.e., a cohort) are defined on the basis of **the presence or absence of exposure** to a suspected risk factor for a disease.
- At the time exposure status is defined, all potential subjects must be free from the disease under investigation
- Eligible participants are then followed over a period of time to assess the occurrence of that outcome.



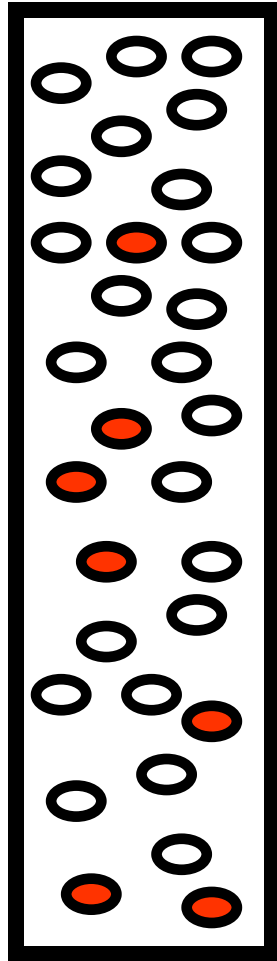
# Cohort studies

## Purpose

- Study if an exposure is associated with outcome(s)?
- Estimate risk of outcome in exposed and unexposed cohort
- Compare risk of outcome in two cohorts



*follow-up period*



*end of follow-up*

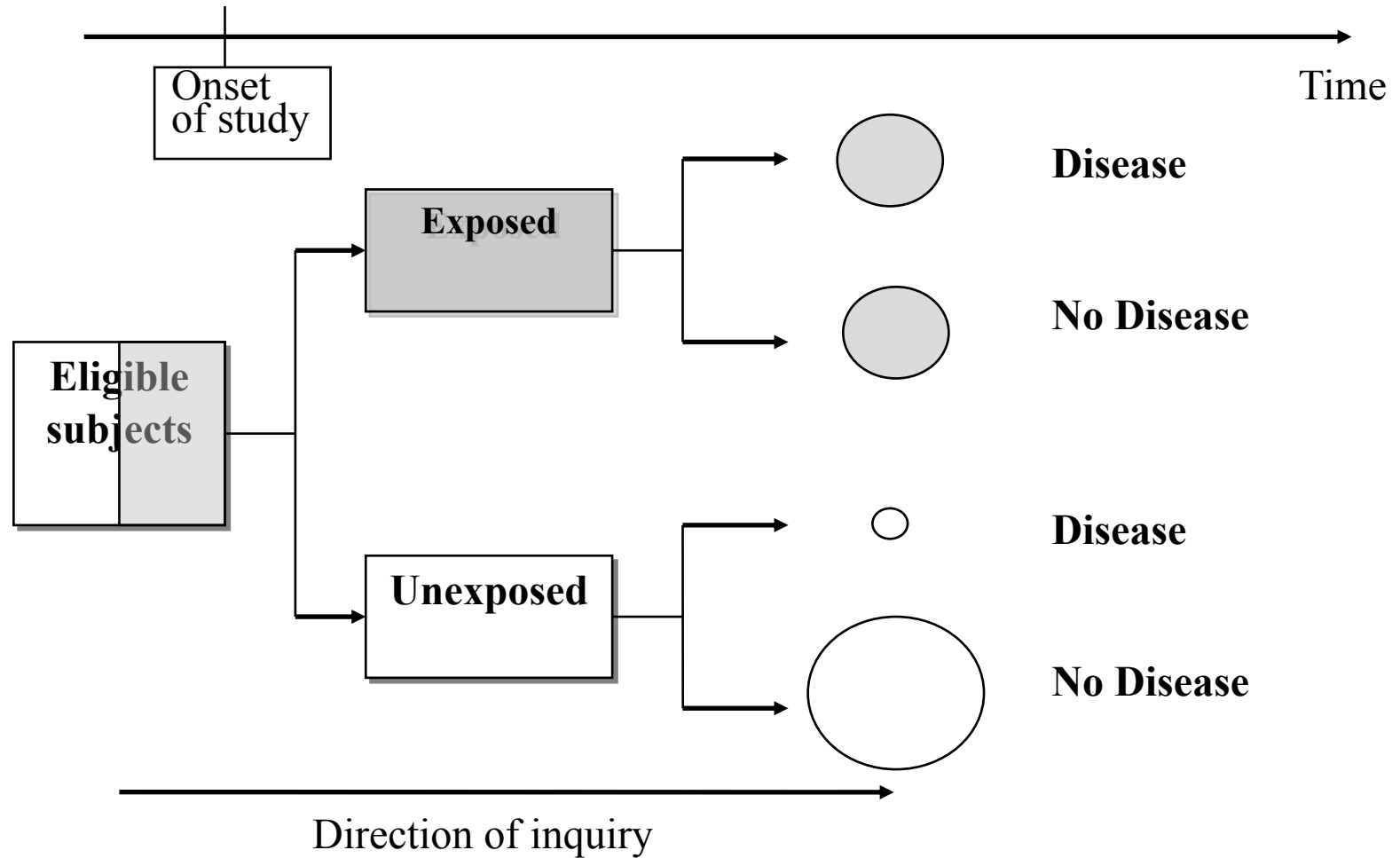
## Calculate

measure of frequency:

- **Cumulative incidence**
  - ☹ Incidence proportion
  - ☹ Attack rate (outbreak)
- **Incidence density**



# Schematic diagram of a cohort study





## Perinatal asphyxia cohort

- In this study, investigators evaluated 49,000 infants whose Apgar scores were recorded at 1 and 5 minutes of age.
- For those infants who did not achieve a score of 8 or higher at 5 minutes, Apgar scores were then recorded at 10, 15, and 20 minutes.
- All the children were then followed to the age of 7 years.

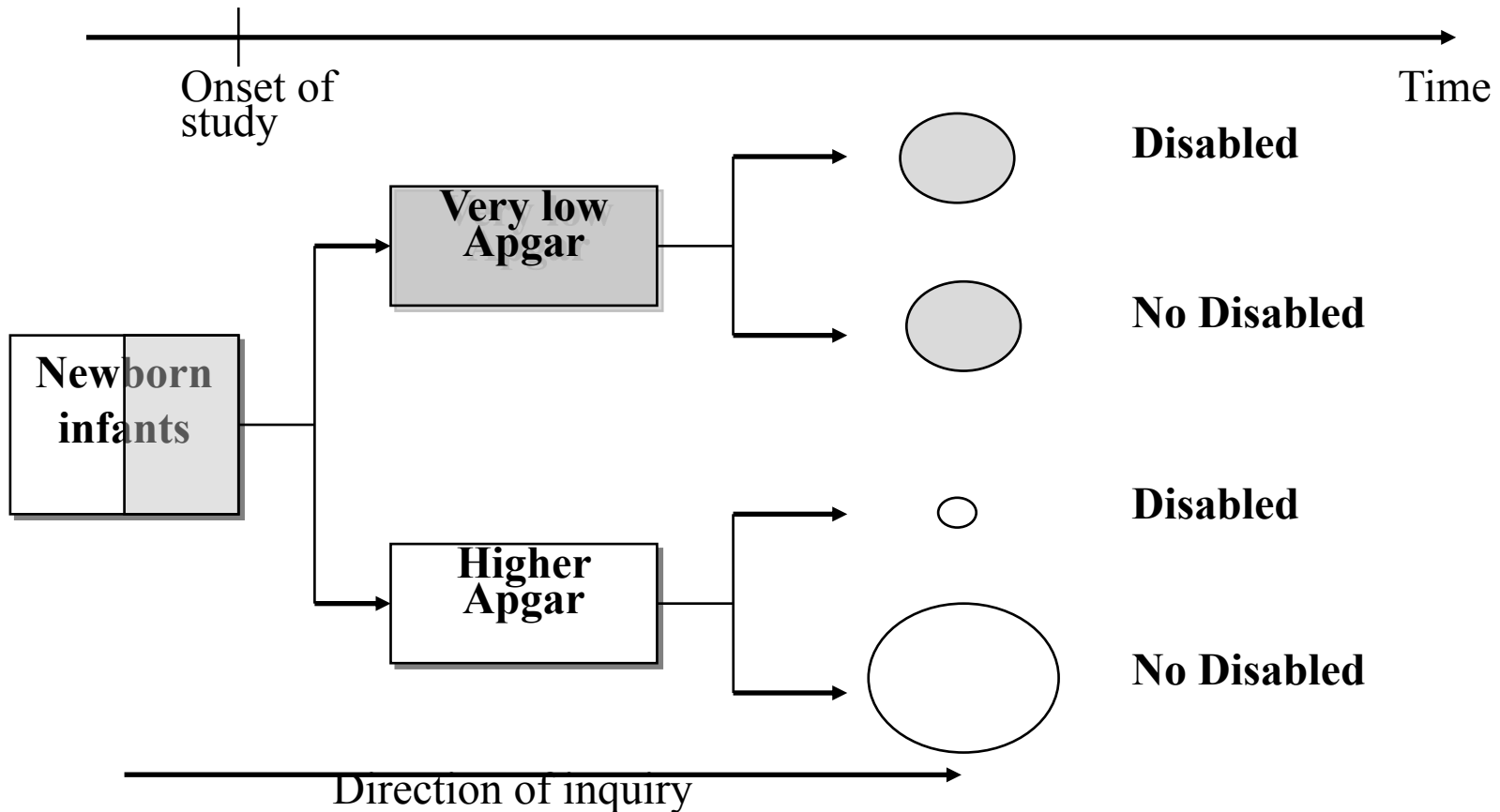


## Follow up of the cohort

- The occurrence of seizures was determined through clinical observations in the newborn nursery;
- Interval histories were recorded at **4, 8, 12, and 18 months** of age and yearly thereafter.
- The presence of cerebral palsy was determined by physical examination at age **7** years.



# A cohort study on the utility of Apgar scores as predictors of chronic neurologic disability





# Outcomes of the cohort study

- This study demonstrated that low Apgar scores are a **risk factor** for the development of cerebral palsy.
- Of the **99** children who survived and had Apgar scores of **0-3** at **10, 15, or 20** minutes, **12** were found to have cerebral palsy. Eleven of those 12 also had delayed mental development.
- However, **55%** of the children with cerebral palsy at age **7** had Apgar scores of **7 or higher** at 1 minute, and 73% scored 7 or higher at 5 minutes.



## The Suggestions of Pediatrician

- Although their baby does have an **increased risk** of cerebral palsy and developmental delay, such an outcome occurs in only about **one of eight** asphyxiated neonates.
- It should be reassuring to the parents to learn that **80%** of even the most severely asphyxiated newborns were free of major neurologic handicap at early school age.



# Examples (1)

<http://www.framinghamheartstudy.org/>

Directions | Contact Info | Search  Go

## FRAMINGHAM HEART STUDY

A Project of the National Heart, Lung and Blood Institute and Boston University

About FHS | Participants | FHS Investigators | Risk Score Profiles | FHS Bibliography | For Researchers

**Three generations of participants.**

*The dedication of our thousands of participants has made, and continues to make, our rigorous epidemiologic research possible.*

**FRAMINGHAM HEART STUDY**  
Three Generations of Dedication  
1948 1971 2002  
A Project of the National Heart, Lung, and Blood Institute and Boston University

### Welcome to the Framingham Heart Study

In 1948, the Framingham Heart Study embarked on an ambitious project in health research to identify the common factors that contribute to cardiovascular disease by following its development over a long period of time in a large group of participants. >>

### THE SABRe in CVD Initiative (Systems Approach to Biomarker Research in Cardiovascular Disease)

For decades, Framingham Heart Study (FHS) participants have generously donated blood samples for research. Hundreds of articles have been published based on analyses of these samples. We store remaining specimens in freezers for future projects. Meanwhile, laboratory techniques have improved, so very small amounts of specimen are enough for a large variety of measurements. In his letter to you, FHS Director, Dr. Daniel

Clinical Epidemiology

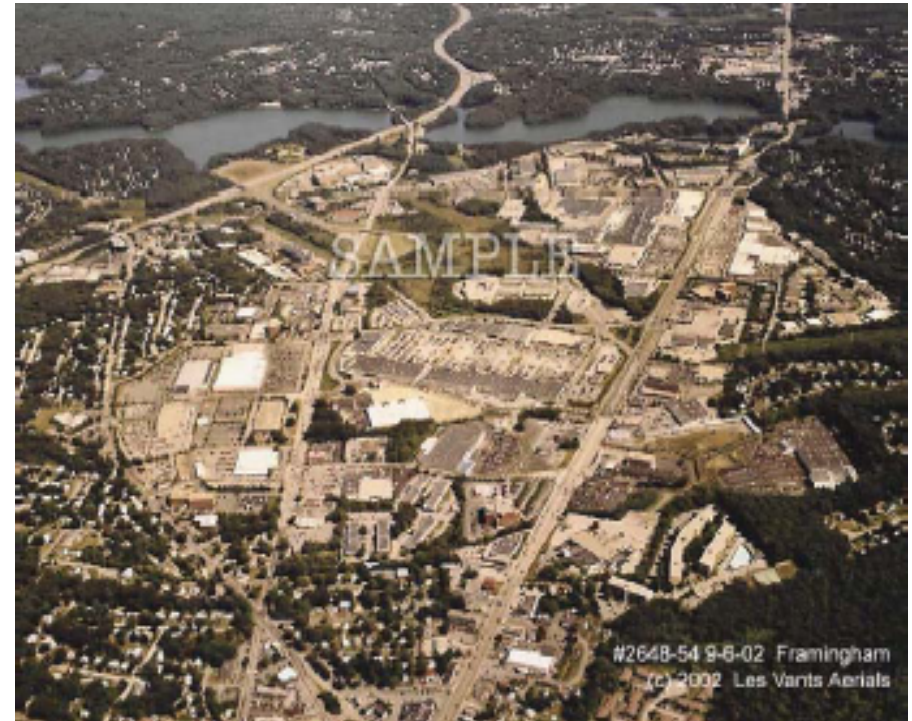


- **Framingham study of cardiovascular disease (CVD)**
  - Individuals 30 – 62 years old in community at risk for disease
  - Framingham, MA, 1948 to present
  - To identify the common factors or characteristics that contribute to CVD



## Why Framingham?

- Framingham participated in a major study of tuberculosis in 1918 .
- The presence of several large employers and the support of a well-informed and highly cooperative and stable medical community.
- Framingham is proximate to Boston's major medical centers.





## Original Cohort Sample

### Original Cohort: Age-Sex Distribution At Entry (1948)

Age	29-39	40-49	50-62	Totals
Men	835	779	722	2336
Women	1042	962	869	2873

### As of February 1998

Age	70-79	80-89	90-99	100 +	Totals
Men	75	243	34	1	353
Women	124	500	114	4	742



## Offspring Study Sample

### Offspring Cohort: Age-Sex Distribution At Entry (1971)

Age	< 10	10-19	20-29	30-39	40-49	50-59	60-70	Totals
Men	--	126	544	789	694	298	38	2489
Women	6	113	692	836	739	246	14	2646

### As of February 1998

Age	30-39	40-49	50-59	60-69	70-79	80-89	90-100	Totals
Men	32	263	766	619	375	32	1	2088
Women	29	347	923	684	422	30	1	2436



# Measurements

## Exposure

- **Content** - Nature of the exposure; biologic mechanisms
- **Quality**
  - Continuous - e.g., serum cholesterol
  - Periodic - e.g., cigarettes, sexual contacts
  - Singular - e.g., nuclear exposure
- **Quantity**
  - Continuous and periodic exposures must be quantified
  - **Dose-response** relationship



## Exposure

- Preexisting records (Occupational exposures, medications)
- Self reported
- Questionnaire
- Lab investigations & physical measurements
- Multiple sources (Framingham Heart study)



# Measurements

## Outcome Definition

- **Primary outcome** - the main event that will be related to the exposure
- **Secondary outcomes** - other events that are of interest and may corroborate the findings of the main outcome



## Outcome

Disease identification should **be comparable** between exposed and unexposed groups and investigators should **be blinded to exposure**

- **Population based disease registries** e.g.. cancer registries
- **Hospital records** : for diseases that require hospitalization
- **Direct examination**: much more standardized than registries and records but expensive



## Follow-up

- All subjects must have an equal likelihood for detecting the outcome
- Disease ascertainment must be **comparable** between the exposed and unexposed subjects
- Follow-up period



## Follow-up

- **Follow-up period**
  - The length of the required period of follow-up, or the interval that **elapses** between definition of exposure status and ascertainment of outcome, will be related to the length of latency period for the outcomes of interest.
  - In general, the longer the observation period required, the more difficult it will be to achieve complete follow-up. (move, changing job, name and so on)



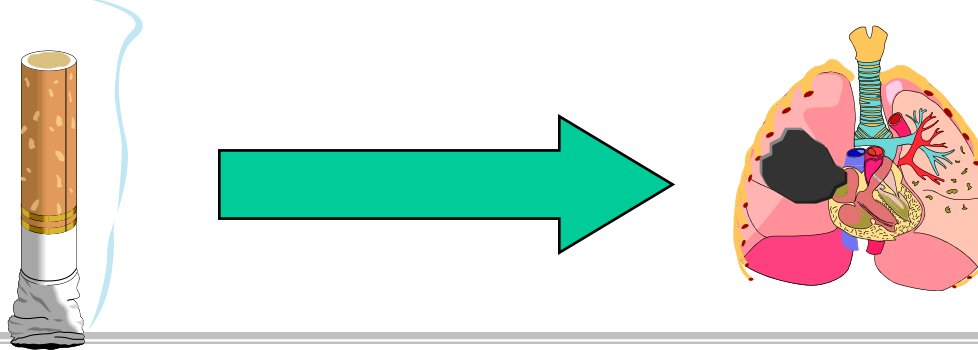
# Limitations

- Its participants are mainly white, middle class individuals.

## Example (2) **Smoking and lung cancer**

### Background

- The mortality of lung cancer was increased rapidly in the early of 20 century.
- Ecological studies showed that the mortality of lung cancer and the tobacco consumption had a parallel relationship.



- In 1948, Doll and Hill conducted a case-control study on the lung cancer.
- In 1951, Doll and Hill started a cohort study.





**Subjects : 59600 doctors**

**Exposure :**

- **Non-smoking ( control group )**
- **Smoking ( 1 cigarette /day ) ( exposure group )**
  - **1 - 14**
  - **15 - 24**
  - **25+**



## Data collection

- Questionnaire on the conditions of common things and smoking
- Finally, 40 701 questionnaires were collected(69.29% )

## Result

The mortality of the subjects in this study was 1 per 100 000 (followed 20 years )

## Conclusion

Smoking was associated with the increased risk of lung cancer, and the more consumption of tobacco, the more risk of lung cancer. (dose-response )

## Example (3)

### Thalidomide and Congenital Malformation

**During 1959-1961 , the incident rate of neonate nanomelia ( like seal with short upper limbs ) was increased obviously in west Europe, especially in west German and U.K. .**

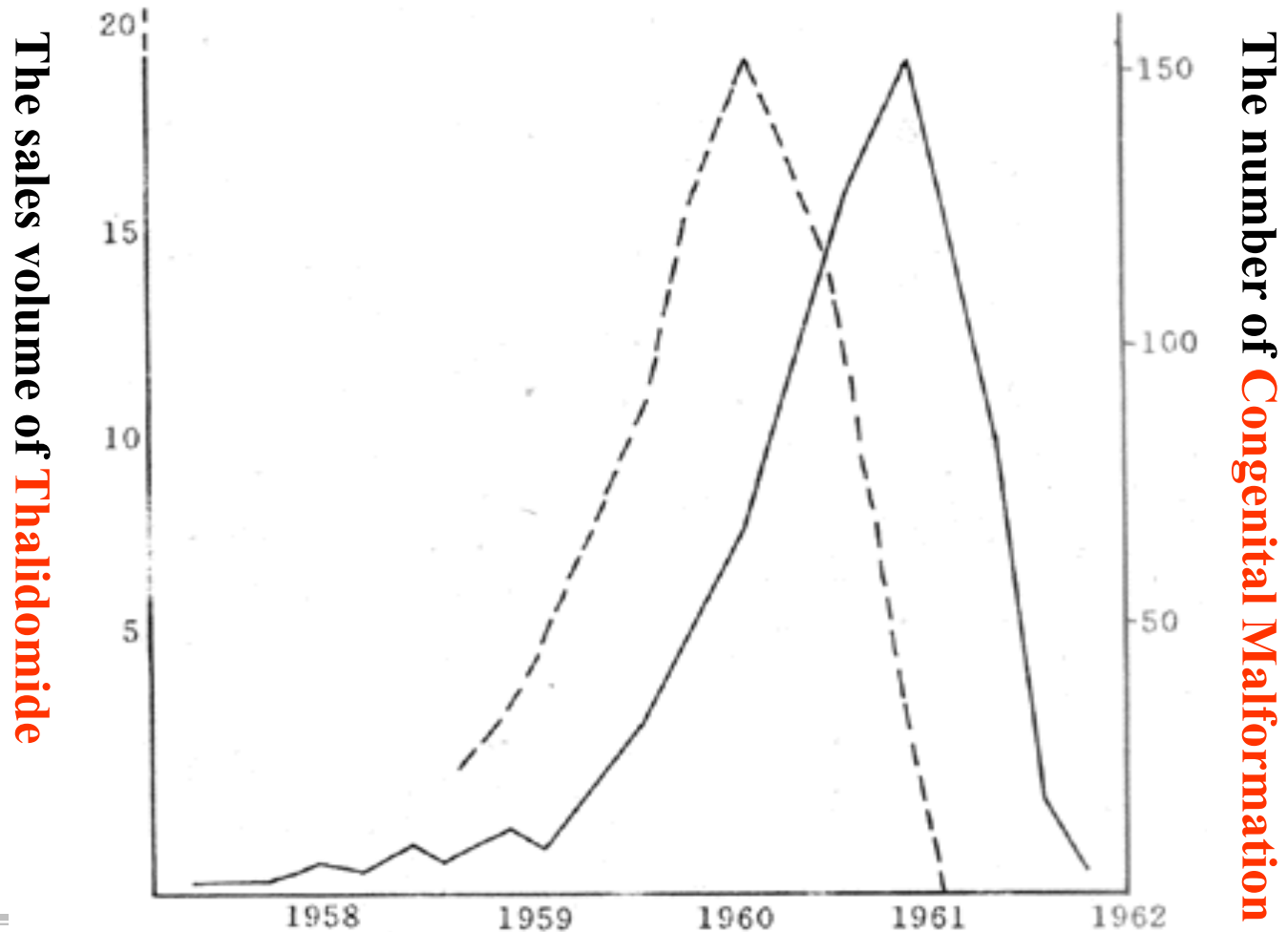
**Nearly more than 10 000 neonate were delivery with short upper limbs malformations, and this was a disaster in the history of medicine.**



Malformations due to maternal ingestion of thalidomide (Schardein 1982 and Moore 1993).



## The sales volume of thalidomide were consistent with the numbers of congenital malformation babies



## Thalidomide baby



**Many children were divested the upper limbs**



## Cohort study on the Congenital Malformation

The relationship between the thalidomide taken by the pregnancy and the Congenital Malformation

Thalidomide Malformation	Exposed Unexposed		Total
	YES	NO	
YES	10	51	61
NO	14	21434	21448
<b>Total</b>	<b>24</b>	<b>21485</b>	<b>21509</b>



***CONCLUSION* : The thalidomide taken by the pregnancy women was associated with significantly increased risk of Congenital Malformation**

□ Incident rate of exposure group =  $10/24 = 42\%$

Incident rate of non-exposure group =  $51/21485 = 0.24\%$

□  $RR = (10/24) \div (51/21485) = 175.95$



# Timing of Measurement

- A cohort study is usually **prospective**, that is, exposure to the risk factor and subsequent health outcomes are observed after the beginning of the study.
- Occasionally, a cohort study is **retrospective** (or historical), that is, it utilizes information on prior exposure to the risk factor and subsequent disease status.

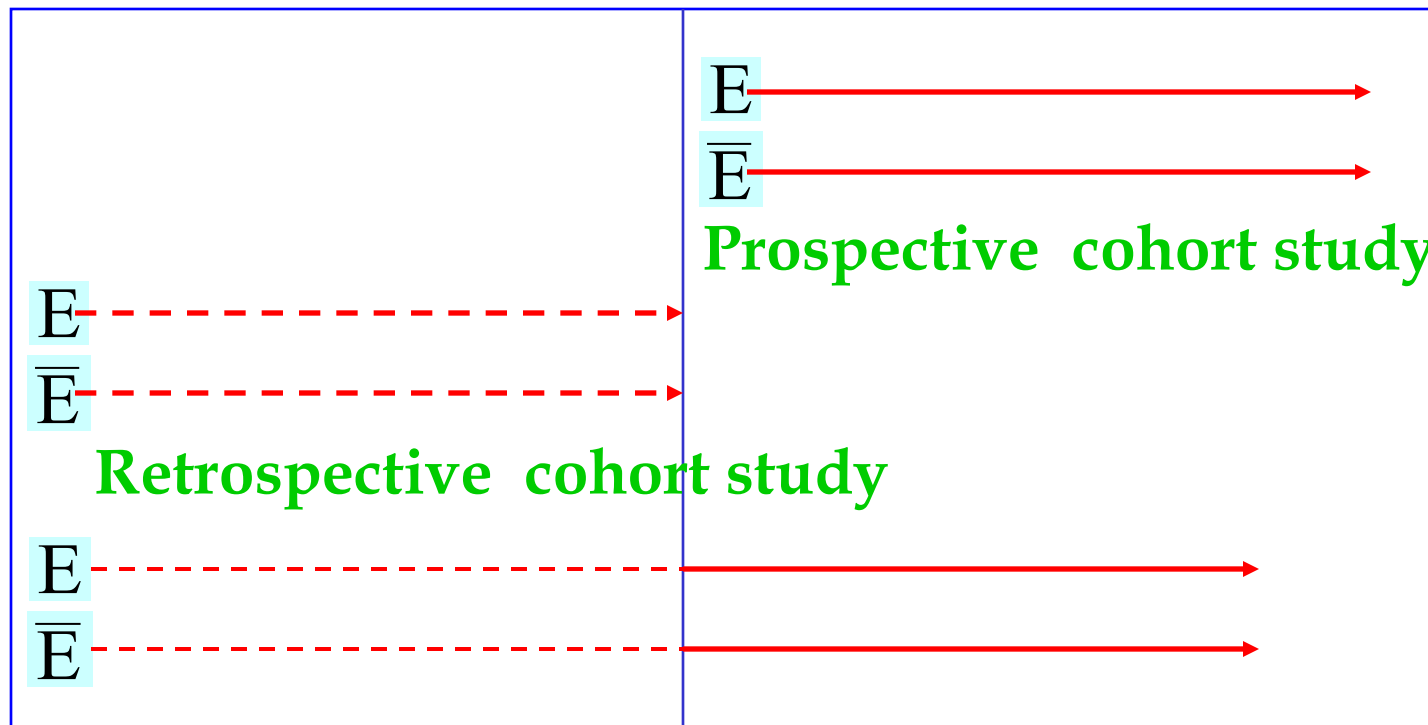


# Types of cohort study

Past

Now

Future



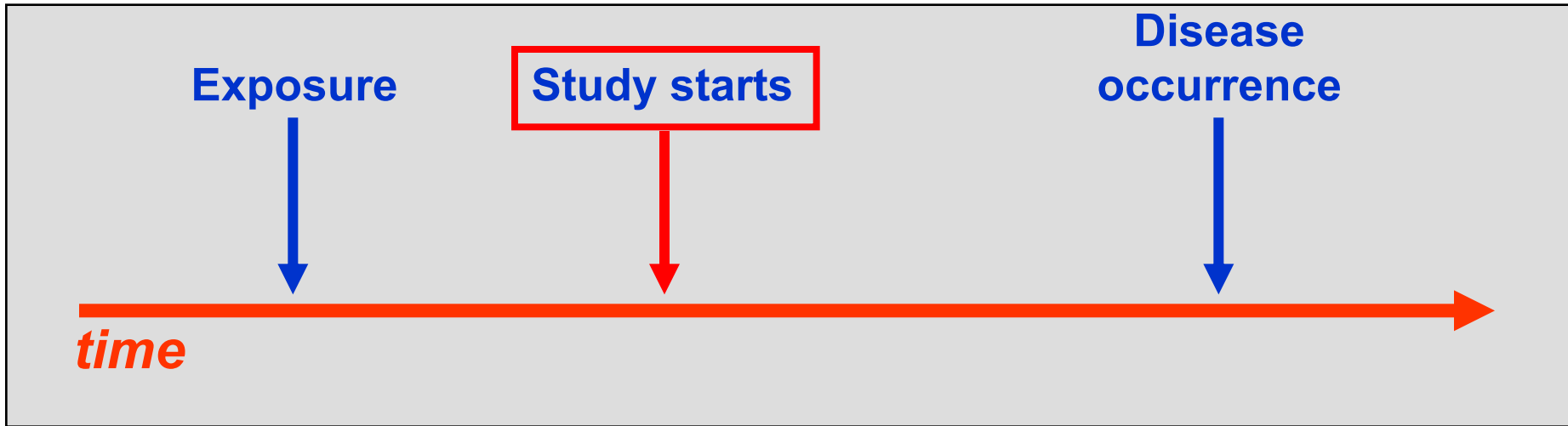
Prospective cohort study

Retrospective cohort study

Ambispective cohort study



# Prospective cohort study



- Identify your cohort in the **present**
- Determine exposure status or possible explanatory/prognostic factors in the **present** or in the **future**
- Ascertainment of outcome done in **future**



## Example:

- ◆ A prospective cohort study of neonatal asphyxia and subsequent mental retardation could be started in 2000.
- ◆ The degree of neonatal asphyxia could be determined for births occurring through 2001, and the development of mental retardation could be assessed between 2001 and 2006, or later.



# Prospective cohort study

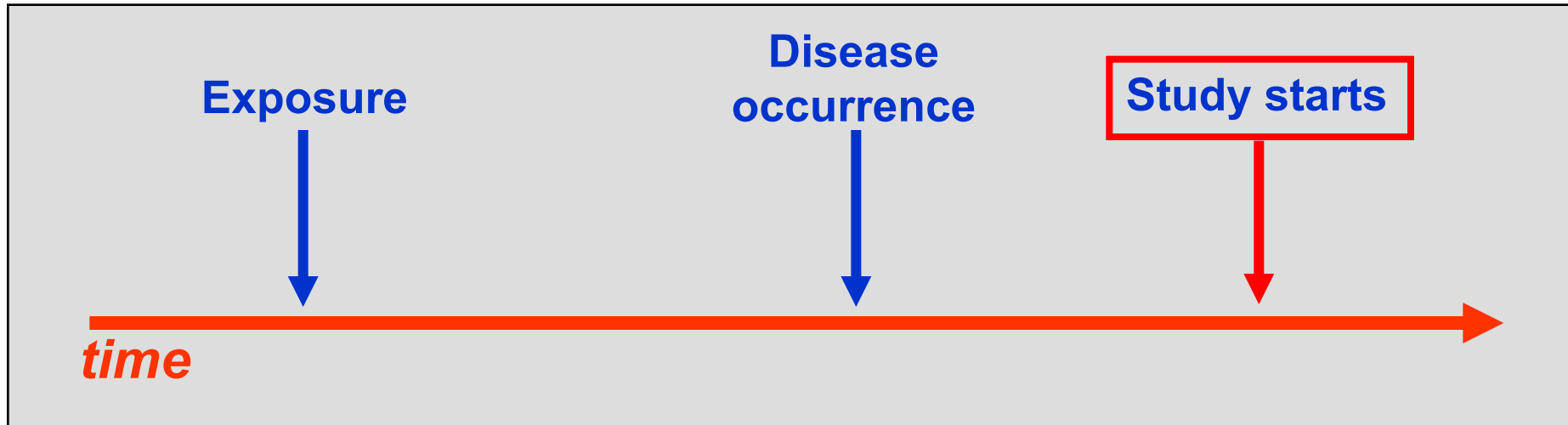
## Strengths:

- **time sequence strengthens inference about cause**
- **more accurate measurement of prognostic or risk factors (do not need to reconstruct past exposures)**
- **more complete measurement of confounding**
- **can study situations where randomization is not possible**

## Weaknesses:

- **expensive and time consuming (inception and follow-up)**
- **large numbers required to study rare outcomes**
- **difficult to study chronic diseases with long latency**

# Retrospective cohort studies



- Identify cohort in the **past** Eg, through records or administrative databases
- Determine exposure or prognostic factors in the **past**  
Again, records or databases
- Outcome can be identified in **past** or **present**
- Outcome must be **after** previous two steps

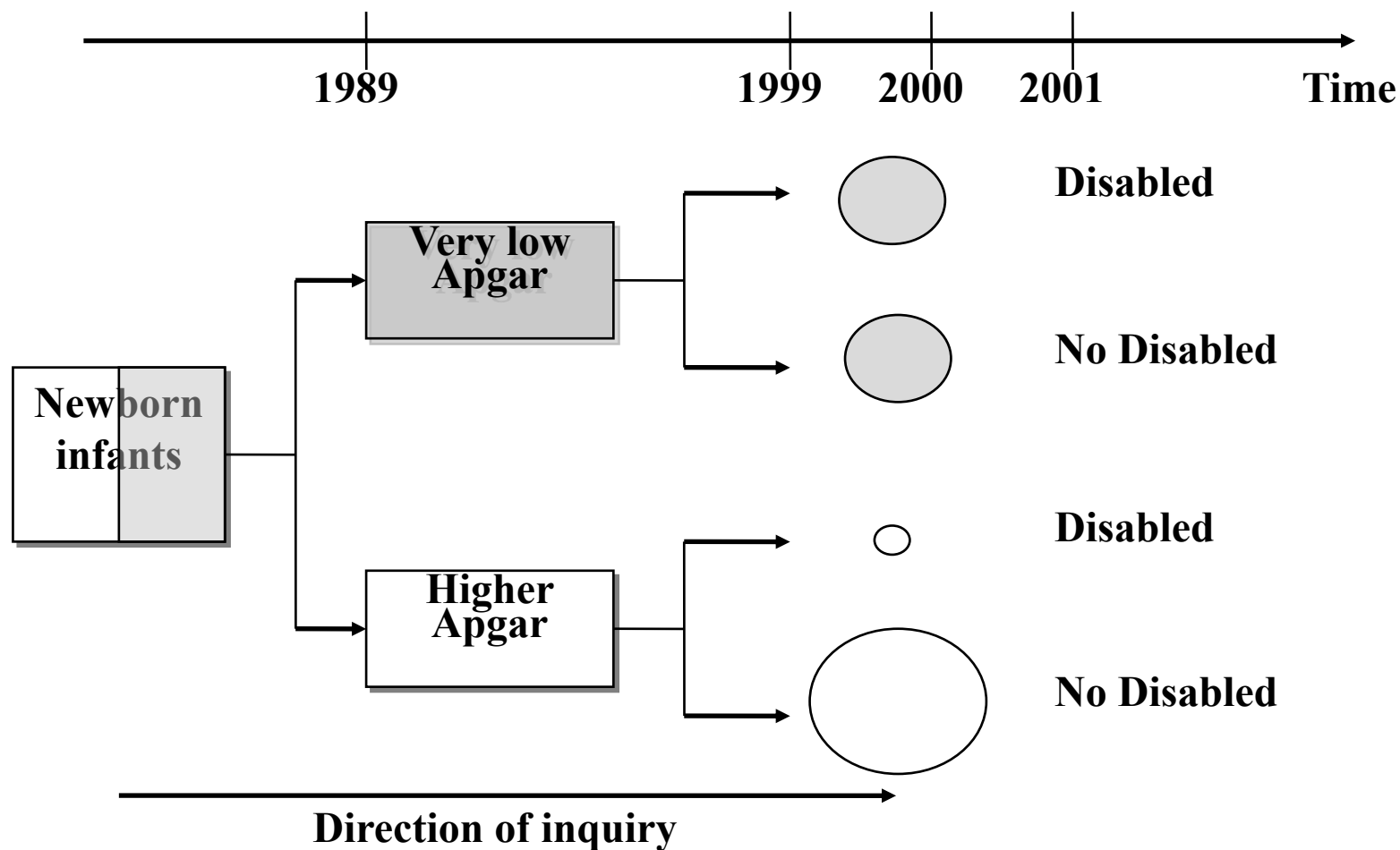


## Example:

- A retrospective cohort study of neonatal asphyxia and neurologic disability designed in 2001 might involve a review of the medical records of infants born in a particular hospital in 1989 to determine the level of asphyxia.
- Followed by a review of school achievement records over the period 1999-2000 to determine the degree of intellectual functioning.
- Note that exposure to risk factors and the subsequent development of the health outcome occur prior to the beginning of the retrospective cohort study.



# Schematic diagram of a retrospective cohort study





# Retrospective cohort studies

## Strengths:

- cohort easier to assemble (inception period in past)
- baseline measurements already available
- follow-up period already taken place
- less costly and time-consuming

## Weaknesses:

- no control over the quality of past measurements
- incomplete data sets
- control for confounding may be incomplete



## Comparison of the attributes of retrospective and prospective cohort study.

<b>Attribute</b>	<b>Retrospective Approach</b>	<b>Prospective Approach</b>
<b>Information</b>	<b>Less complete and accurate</b>	<b>More complete and accurate</b>
<b>Discontinued exposures</b>	<b>Useful</b>	<b>Not useful</b>
<b>Emerging new exposures</b>	<b>Not useful</b>	<b>Useful</b>
<b>Expense</b>	<b>Less costly</b>	<b>More costly</b>
<b>Completion time</b>	<b>Shorter</b>	<b>Longer</b>



# Advantages and disadvantages of cohort studies

Advantages	Disadvantages
Direct calculation of risk ratio (relative risk)	Time consuming
May yield information on the incidence of disease	Often require a large sample size
Clear temporal relationship between exposure and disease	Expensive
Particularly efficient for study of rare exposures	Not efficient for the study of rare diseases
Can yield information on multiple exposures	Losses to follow-up may diminish validity
Can yield information on multiple outcomes of a particular exposure	Changes over time in diagnostic methods may lead to biased results
Minimizes bias	
Strongest observational design for establishing cause and effect relationship	



# Selection of Subjects

- ◆ **The type of exposure under investigation.**
- ◆ **The frequency of the exposure in the population.**
- ◆ **The accessibility of subjects and the likelihood of their continuing participation.**



- Both exposed and unexposed groups must **be free of** the outcome of interest at the start of the study.
- They must be **similarly eligible** to develop the outcome of interest during the course of the study.



## Selection of Exposed Group

- Depends on research question
- Depends on frequency of exposure
  - **Common exposures: general population**
  - **Rare exposures: selected groups**
    - E.g., selected groups used in rare occupational exposures to chemicals
    - E.g., radiation and ankylosing spondylitis
- Outcome must not be rare in exposed
  - **Attributable risk must be high: e.g., if the disease is rare, even among the exposed, cohort study is inefficient.**
- Accessible and compliant subjects
  - E.g., Nurse's Health Study, Health maintenance organizations, residents of stable communities, and labor union members.



# Selection of Exposed Group

(will depend on the specific hypothesis)

## From the general population

- Framingham Heart study - still confined to over 30 yrs subjects (high risk group)

## Specific populations

- British doctors study (1950s)  
cooperation, easy to follow up, accuracy on medical information

## Occupational exposures

- Asbestos workers to study the effects of asbestos
- Radiologists to study the effects of radiation  
rare exposures, availability of records, initial and periodic medical examination, easy to follow up



- The degree of exposure depending on the goals of study:

e.g., exposed or unexposed.

a range of exposure levels:

**Dichotomous**

Apgar score 0~3

Apgar score >3(cutoff point)

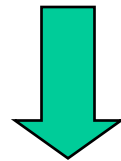
**Multiple ordered**

0~3

4~6

7~10

**Continuous-by gradations**



Whether a relationship exists between the **dose** (of the exposure) and the **response**.



## Selection of Unexposed Cohort

### “Comparison” or “Control” group

- Similar as possible to exposed
  - Control for other factors, other than the exposure: e.g., use non-exposed family members, other workers from similar occupation, but not exposed.
- General Population?
  - Sometimes OK, but not as healthy as workers, therefore not good for comparisons. ?
- Multiple groups
  - E.g., asbestos exposure and mesothelioma: compare to cotton textile workers and also the general population.



## Selection of Comparison Cohort

- ◆ In the study of perinatal asphyxia: the unexposed group was defined as the infants with the highest Apgar scores (7-10), indicating the lowest degree of perinatal asphyxia.
- ◆ With randomization, factors known to be related to the development of disease—as well as other factors not yet recognized as related to the disease—tend to be **balanced** between the groups.



## Guidelines for selection of exposed and unexposed subjects in cohort studies.

Unexposed	Exposed	Unexposed and Exposed
Both exposed and unexposed groups	The baseline characteristics of exposed persons should not differ systematically from those of unexposed persons, except for the exposure of interest	Unexposed persons should be sampled from the same (or comparable) source population as the exposed group
should be free of the disease of interest and equally susceptible to development of the disease at the beginning of the study		Multiple comparison groups of unexposed subjects chosen in different ways may reinforce the validity of findings
Equivalent information (quantity and quality) should be available on exposure and disease status in the exposed and unexposed groups		
Both groups should be accessible and available for follow-up		



# DATA COLLECTION

- Exposure (independent variable)
- Clinical response (dependent variable)



# Exposure

## Measurements of exposure used in cohort studies.

Measurements of Exposure	Examples
Intensity	Mean blood pressure level
Duration	Weeks of hypertension
Regularity	Number of affected pregnancies
Variability	Range of measured blood pressures

Example: gestational hypertension



# Clinical response

The degree of surveillance for disease should be similar in the exposed and unexposed groups.

- ◆ a short follow-up period :the perinatal asphyxia and death within the first week of life.
- ◆ a long follow-up period: the perinatal asphyxia and chronic neurologic disability



- Information on outcome status come from **various sources**: records of physicians and hospitals or collecting information .
- **Appropriate diagnostic test**: the accuracy and reliability of diagnosis must not differ between the groups. (standard neurologic examination and psychological tests )



# Analysis

## Risk Ratio (RR)

Outcome <sup>a</sup>	Exposed	Unexposed	Total
Death	<b>A</b>	<b>B</b>	<b>A+B</b>
No death	<b>C</b>	<b>D</b>	<b>C+D</b>
Total	<b>A+C</b>	<b>B+D</b>	<b>A+B+C+D</b>

- (A) Exposed persons who later die
- (B) Unexposed persons who later die
- (C) Exposed persons who do not die
- (D) Unexposed persons who do not die



## Risk Ratio

- Among exposed persons, the risk ( $R$ ) defined as

$$\begin{aligned} R_{(\text{exposed})} &= \frac{\text{Exposed persons who die}}{\text{All exposed persons}} \\ &= \frac{A}{A + C} \end{aligned}$$

- Among unexposed persons, the risk of death is defined as

$$\begin{aligned} R_{(\text{unexposed})} &= \frac{\text{Unexposed persons who die}}{\text{All unexposed persons}} \\ &= \frac{B}{B + D} \end{aligned}$$



- One approach to contrasting the risk in two groups is to create a ratio measure. The **risk ratio (RR)** or **relative risk** is

$$RR = \frac{R(\text{exposed})}{R(\text{unexposed})} = \frac{A / A + C}{B / B + D}$$



- $RR=1$

The exposed and unexposed persons have the same risk of death and exposure is not related to the outcome. (ie, the null value)

- $RR>1$

The risk among exposed persons is greater than the corresponding risk among unexposed persons, (ie, hazardous exposure)

- $RR<1$

The risk among exposed persons is smaller than the corresponding risk among unexposed persons (ie, beneficial exposure)



# Example

- Relationship between 10 minute Apgar scores and risk of death in the first year of life among children with birth weights of at least 2500 g

	<b>Apgar Score 0-3</b>	<b>Apgar Score 4-6</b>	<b>Total</b>
<b>Death</b>	<b>42</b>	<b>43</b>	<b>85</b>
<b>No death</b>	<b>80</b>	<b>302</b>	<b>382</b>
<b>Total</b>	<b>122</b>	<b>345</b>	<b>467</b>



- The risk among exposed newborns is

$$R_{(\text{exposed})} = \frac{42}{122} = 0.344 = 34.4\%$$

- The risk among “less exposed” newborns is

$$R_{(\text{less exposed})} = \frac{43}{345} = 0.125 = 12.5\%$$

- Quantification of the magnitude of this effect is achieved by calculating the risk ratio :

$$RR = \frac{42}{122} / \frac{43}{345} = 2.8$$



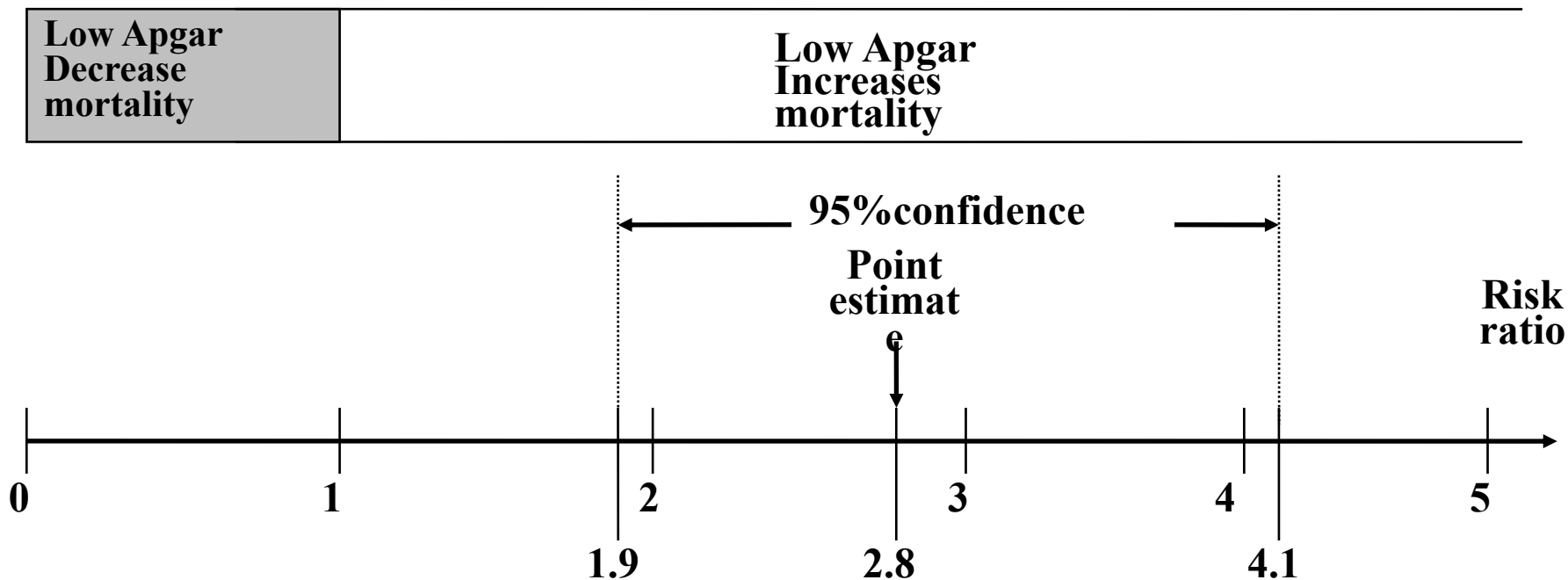
- **RR means:**

**The *RR* is a measure of the strength of association between exposure and outcome.**

**The *RR* of 2.8 means that newborns at this birth weight with 0 ~3 Apgar scores are almost **2.8** times **more** likely to die in the first year of life **than** similar-weighting newborns with 4~6 10-minute Apgar scores.**



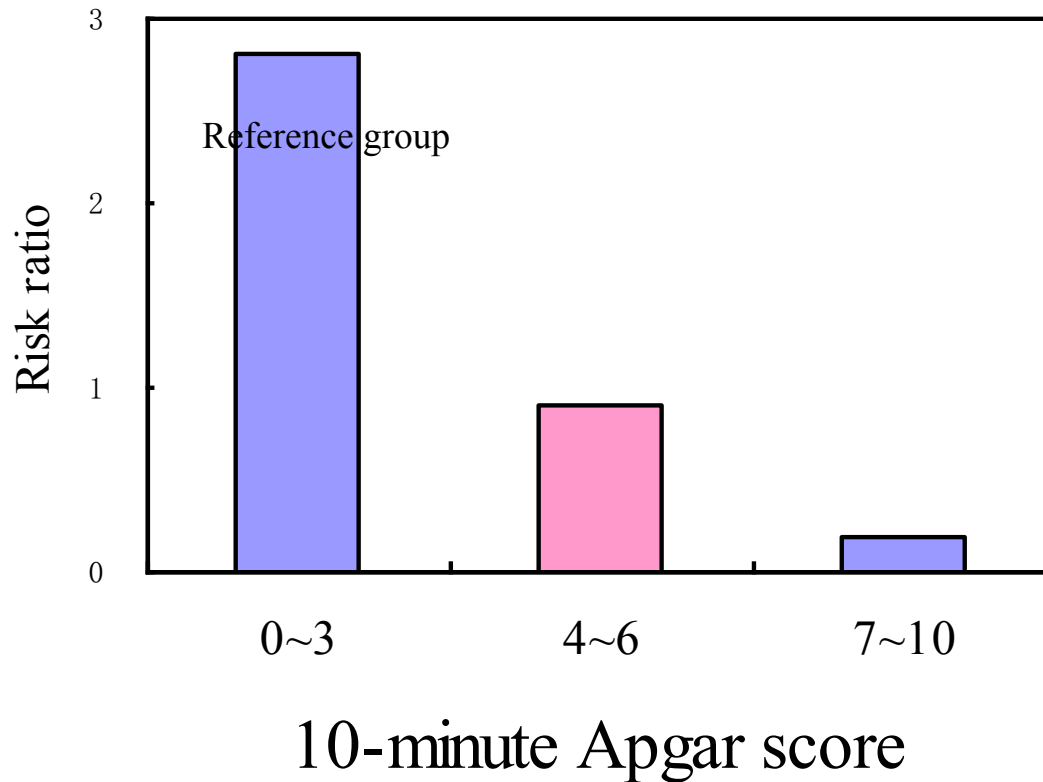
# Point estimate and 95% confidence interval for risk ratio





- Comparison of the risk of death in the group with Apgar scores of 4-6, a third group of newborns, with Apgar scores of 7-10, yields a risk ratio of 0.15, with an approximate 95% confidence interval of (0.11, 0.21).
- This result means that newborns with a 10-minute Apgar score of 7-10 have only about one-sixth the risk of death in the first year of life as newborns with Apgar scores of 4-6.

## **Dose-response** relationship for the association between 10-minute Apgar scores and risk of death among newborns.



**A clear trend of decreasing risk ratio with increasing Apgar score is seen.**



# Attributable Risk Percent

- “risk difference, or excess risk”, or "attributable risk"
- Means: the risk for, one group can be subtracted from the risk for another group .
- The risk difference (**RD**) is defined as

$$\begin{aligned}RD &= R_{(\text{exposed})} - R_{(\text{unexposed})} \\ &= \frac{A}{A + C} - \frac{B}{B + D}\end{aligned}$$



- For the Apgar score-infant mortality data (0~3 versus 4-6), the risk difference (RD) is:

$$RD = \frac{42}{122} - \frac{43}{345} = 0.344 - 0.125 = 0.219$$



- That is, the risk of death in the first year of life is increased by **0.219** for newborns who weigh more than 2500 g and have a 10-minute Apgar score of **0-3**, compared with similar-weighting newborns with a 10-minute Apgar score of **4-6**.



# The attributable risk percent (*ARP*)

- in which the **risk difference** is expressed as a **percentage** of the total risk experienced by the exposed group:

$$\begin{aligned} ARP &= \frac{R_{(\text{exposed})} - R_{(\text{unexposed})}}{R_{(\text{exposed})}} \times 100 \\ &= \frac{A/A + C - B/B + D}{A/A + C} \times 100 \end{aligned}$$



- For the Apgar score-infant mortality data, the attributable risk percent is:

$$ARP = \frac{(0.344 - 0.125)}{0.344} \times 100 = 63.7\%$$

- In other words, almost **two thirds** of the total risk of infant mortality for newborns who weigh more than 2500 g and have 10-minute Apgar scores of 0-3 is related to an Apgar score below the 4-6 level.



- The attributable risk percent typically is used as an indicator of the **public health impact** of exposure.
- These data suggest that birth asphyxia is a **major contributor** to but not the **sole cause** of--infant mortality among severely asphyxiated children.



## Rate Ratio

- In a cohort study, the measured outcome may be an **incidence (or mortality) rate** rather than a risk.
- Summary format of rate data from a cohort study:

	<b>Exposed Persons</b>	<b>Unexposed Persons</b>	<b>Total</b>
<b>Number of outcomes</b>	<i>A</i>	<i>B</i>	<i>A+B</i>
<b>Person-time (PT)</b>	<i>PT</i> (exposed)	<i>PT</i> (unexposed)	<i>PT</i> (total)



**The rate ratio is derived as follows:**

$$\begin{aligned} \text{Rate ratio} &= \frac{\text{Rate of outcome among exposed persons}}{\text{Rate of outcome among unexposed persons}} \\ &= \frac{A/PT(\text{exposed})}{B/PT(\text{unexposed})} \end{aligned}$$



- The magnitude of the rate ratio is interpreted in the same manner as the risk ratio:
  - $RR < 1$  : protective effect,
  - $RR = 1$  : no effect,
  - $RR > 1$  : harmful or risk effect of exposure

The farther away from the null value, the stronger the association between exposure and the rate of the outcome.



## For Example

- the Chicago Heart Association Detection Project in Industry (Dyer et al, 1992)
  - ◆ Subjects :40,000 men and women at 84 cooperating companies.
  - ◆ Object: screened for risk factors for cardiovascular disease and followed an average of 14-15 years.



## Relationship between baseline serum cholesterol level and subsequent mortality rate from coronary heart disease.

	Cholesterol Level		Total
	5.2-6.2mmol/L <sup>b</sup>	≤5.1 mmol/L <sup>c</sup>	
Deaths	26	14	40
Person-years	36,581	68,239	104,820



- The rate ratio is

$$\text{Rate ratio} = \frac{26/36,581}{14/68,239} = 3.5$$

**Conclusion:** The mortality rate from CHD among white males with borderline *high* cholesterol levels was about **3.5** times **higher than** that of white males with *lower* cholesterol levels.

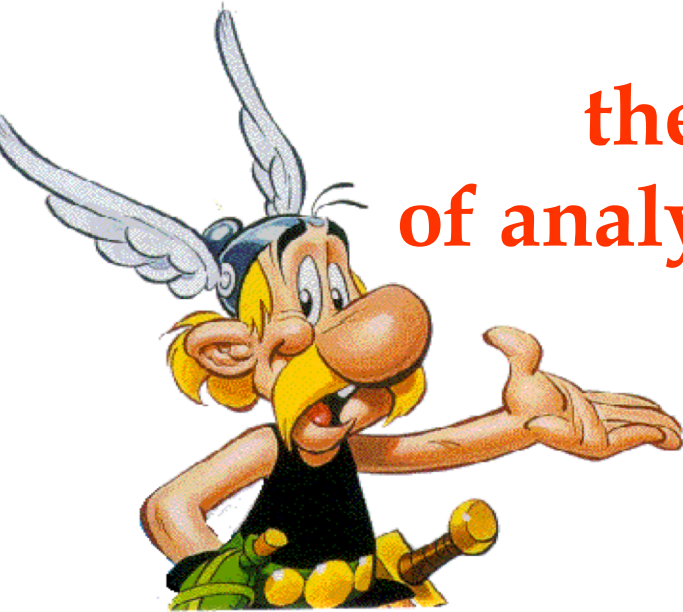


- Adjustment for underlying age differences in the study groups reduced the observed rate ratio to **3.1**.
- Compared with the white males with normal serum cholesterol levels ( $<5.1$  mmol/L), the white males with high serum cholesterol levels ( $>6.2$  mmol/L) , had a **4.1** time increased risk of mortality from CHD.
- A **dose-response relationship** was evident between baseline serum cholesterol level and subsequent CHD mortality.

## Summary

The cohort study  
is  
the gold-standard  
of analytical epidemiology.

Alain Moren





## Summary

- Definition of the cohort study.
- Prospective and retrospective cohort studies
- Measure of effect:

**risk ratio**

**risk difference**

**attributable risk percent**

**rate ratio**



# Checklist for the evaluation of published cohort studies.

## Hypothesis

- A. Is the study hypothesis clearly stated?
- B. Does it address a question of clinical interest and importance?

## Design

- A. Is the cohort design appropriate for the question to be answered?
- B. Is it feasible to perform a cohort study?

## Study Population

- A. Will the study yield a fair comparison between the exposed and unexposed subjects?
- B. Is the sample size adequate to answer the question of interest?
- C. Do the exposed and unexposed subjects come from the same or different population?
- D. Are the exposed and unexposed subjects examined concurrently?
- E. Does the investigator present a rationale for the choice of study population?
- F. Is the study population similar to the type seen in clinical practice?

**Exposure**

A. Has the exposure been defined?

B. What is the source of exposure information?

C. Has the exposure been measured appropriately?

D. Are there objective measures or markers to substantiate subjective measures?

E. Is the exposure an acute or chronic one?

F. For chronic exposures, is there remeasurement during the course of the study?

G. Is it possible to examine a dose-response relationship?

**Disease**

A. Is the disease clearly defined?

B. What is the source of information about the disease?

C. Is there pathological or other confirmation of disease?

D. Has the presence of disease been assessed in a similar fashion for the exposed and unexposed groups?

E. Were those who assessed disease status blind to subject exposure status?



**Follow-up**

A. Was the period of follow-up adequate for the development of disease?

B. Were appropriate measures taken to maintain subjects in the study?

C. Is there discussion of losses to follow-up?

**Analysis**

A. Was an appropriate analysis performed?

B. Are the results statistically significant?

C. Are the results clinically meaningful?



<http://www.ncbi.nlm.nih.gov/pubmed/>



Search: PubMed

Advanced search Help

cohort study

Search

Clear

### Welcome to PubMed

PubMed comprises more than 19 million citations for biomedical articles from MEDLINE and life science journals. Citations may include links to full-text articles from PubMed Central or publisher web sites.

**Search for whatever you want to know about the medicine.**

### Using PubMed

[PubMed Quick Start](#)

[New and Noteworthy](#)

[PubMed Tutorials](#)

[Full Text Articles](#)

### PubMed Tools

[Single Citation Matcher](#)

[Batch Citation Matcher](#)

[Clinical Queries](#)

[Topic-Specific Queries](#)

### More Resources

[MeSH Database](#)

[Journals Database](#)

[Clinical Trials](#)

[E-Utilities](#)



Search: PubMed

RSS Save search Advanced search Help

cohort study Search Clear

Display Settings: Summary, 20 per page, Sorted by Recently Added

Send to:

Filter your results:

All (767733)

Review (28206)

Free Full Text (121622)

Manage Filters

Results: 1 to 20 of 767733

<< First < Prev Page 1 Next > Last >>

Also try:

- prospective cohort study
retrospective cohort study
birth cohort study
millennium cohort study
multicenter aids cohort study

Titles with your search terms

- Cytochrome P450 2C19 polymorphism in young patients treated with clopidogrel a [Lancet. 2009]
Risk of myocardial infarction and death

- Delayed gadolinium-enhanced magnetic resonance imaging (dGEMRIC) of hip joint cartilage in femoroacetabular impingement (FAI): Are pre- and postcontrast imaging both necessary?
Analysis of breast cancer related gene expression using natural splines and the Cox proportional hazard model to identify prognostic associations.
Reliability and validity of the Dutch version of the maternal antenatal attachment scale.

Thanks for your attention!

