

# **Common Research Designs Used in Epidemiology**

## Common Research Designs Used in Epidemiology

### ■ FUNCTIONS OF RESEARCH DESIGN

### ■ TYPES OF RESEARCH DESIGN

#### ■ Observational Designs for Generating Hypotheses

#### ■ Observational Designs for Generating or Testing Hypotheses

#### ■ Techniques for Data Summary, Cost-Effectiveness Analysis

# FUNCTIONS OF RESEARCH DESIGN

- **Research** is the process of answering a question that can be answered by appropriately collected data.
- All research, whether quantitative or qualitative research, is descriptive, and no research is better than the **quality of the data** obtained.
- The **rules** that govern the **process of collecting and arranging the data** for analysis are called **research designs**.

# Hypothesis Generation

- The process of developing a list of possible candidates for the "causes" of the disease and obtaining initial evidence that supports one or more of these candidates.

# Hypothesis Testing

- When one or more hypotheses are generated, they must be tested by making predictions from the hypotheses and examining new data to determine if the predictions are correct.

# Research Designs

- Some specific research designs are appropriate for **hypothesis generation** and some are appropriate for **hypothesis testing**.
- Some designs can be used for either, depending on the circumstances.
- **No research design is perfect**, however, because each has its advantages and disadvantages.

# Functions of Research Designs

A good epidemiologic research design should perform the following functions:

1. Enable a **comparison** of a variable (e.g., disease frequency) between two or more groups at one point in time or, in some cases, between one group before and after receiving an intervention or being exposed to a risk factor
2. Allow the comparison to be quantified in **absolute** terms (as with a risk difference or rate difference) or in **relative** terms (as with a relative risk or odds ratio)
3. Permit the investigators to determine when the risk factor and the disease occurred, to determine the **temporal** sequence
4. **Minimize biases, confounding,** and other problems that would complicate interpretation of the data.

# Research Designs

- **Cross-sectional** surveys and **ecologic studies** are useful for developing hypotheses;
- **Cohort studies** and **case-control** studies can be used to develop hypotheses and to test them, although the hypothesis development and hypothesis testing must always be done on different data sets
- **Randomized clinical trials (RCT)** or field trials are usually the best for testing new treatments or preventive measures.



# TYPES OF RESEARCH DESIGN

- Some research questions can be answered by more than one type of research design, the choice of design depends on a variety of considerations, including speed, cost, and availability of data.
- Each type of research design has advantages and disadvantages

# TYPES OF RESEARCH DESIGN

## **Cross-sectional (Prevalence study)**

- **Advantages**

1. Quick and easy to perform
2. Useful for hypothesis generation
3. Useful for measuring current health status

- **Disadvantages**

1. Do not offer evidence of a temporal relationship
2. Subject to **Neyman** (late look) bias
3. Not good for hypothesis testing

# TYPES OF RESEARCH DESIGN

## **Cohort studies** (Incidence study)

- **Advantages**

1. Can be performed retrospectively or prospectively
2. Can be used to obtain a true (absolute) measure of risk
3. Can study many disease outcomes
4. Good for studying rare risk factors

- **Disadvantages**

1. Time-consuming
2. Costly (especially prospective studies)
3. Can study only the risk factors measured at the beginning
4. Can be used only for common diseases
5. Have losses to follow-up

# TYPES OF RESEARCH DESIGN

## Case-control

- **Advantages**

1. Quick and easy to perform
2. Can study many risk factors
3. Good for studying rare diseases

- **Disadvantages**

1. Can obtain only a relative measure of risk
2. Subject to recall bias
3. Selection of controls may be difficult
4. Temporal relationships may be unclear
5. Can study only one disease outcome at a time

# TYPES OF RESEARCH DESIGN

## Randomized controlled trials (RCT)

- **Advantages**

1. "Gold standard" for evaluating treatment interventions (clinical trials) or preventive interventions (field trials)
2. Allow investigator to have extensive control over research process

- **Disadvantages**

1. Time-consuming and usually costly
2. Study only controlled interventions or exposures
3. Problems related to therapy changes and dropouts
4. May be limited in generalizability
5. Often unethical to perform at all

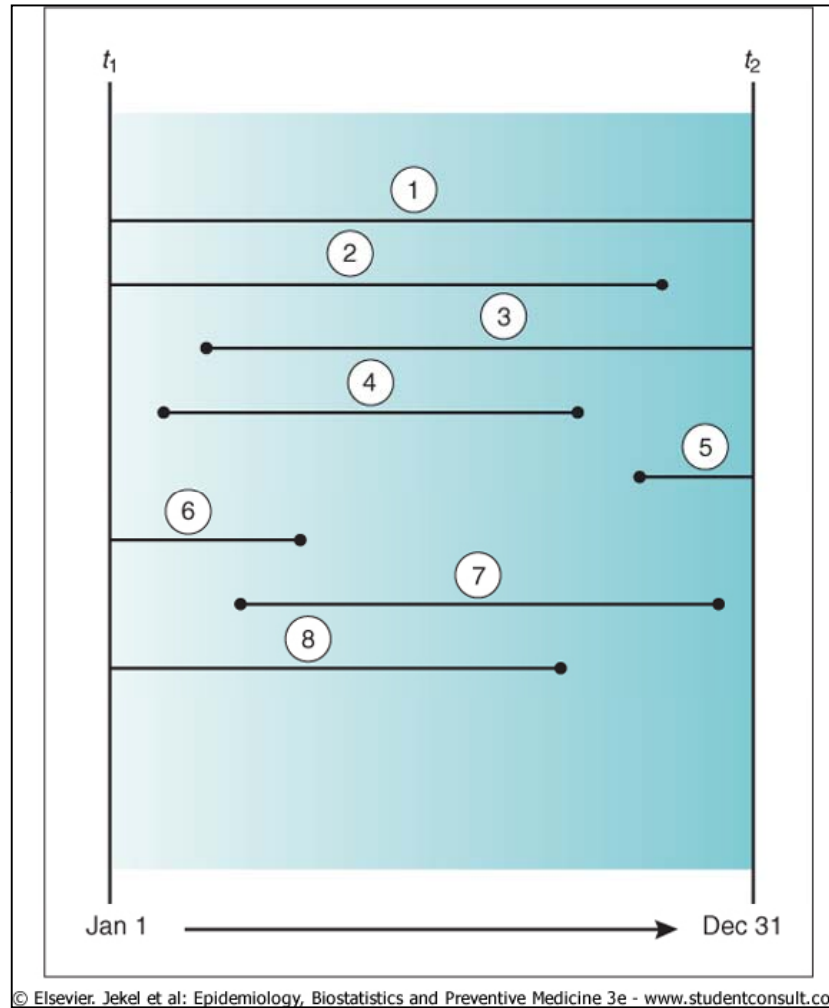
# Observational Designs for Generating Hypotheses

## Cross-Sectional Surveys

A cross-sectional survey is a survey of a population at a single point in time.

Examples are:

- **Interview survey** (interviewers, Tel, mail or Email questionnaire)  
**Neyman or late look bias:** Selects only longer lasting diseases
- **Screening program**  
**Length bias:** Tend to find only less aggressive cases



The longer a line, which represents the duration of a disease, the more likely it is to be touched by any random vertical line (representing a cross-sectional survey).

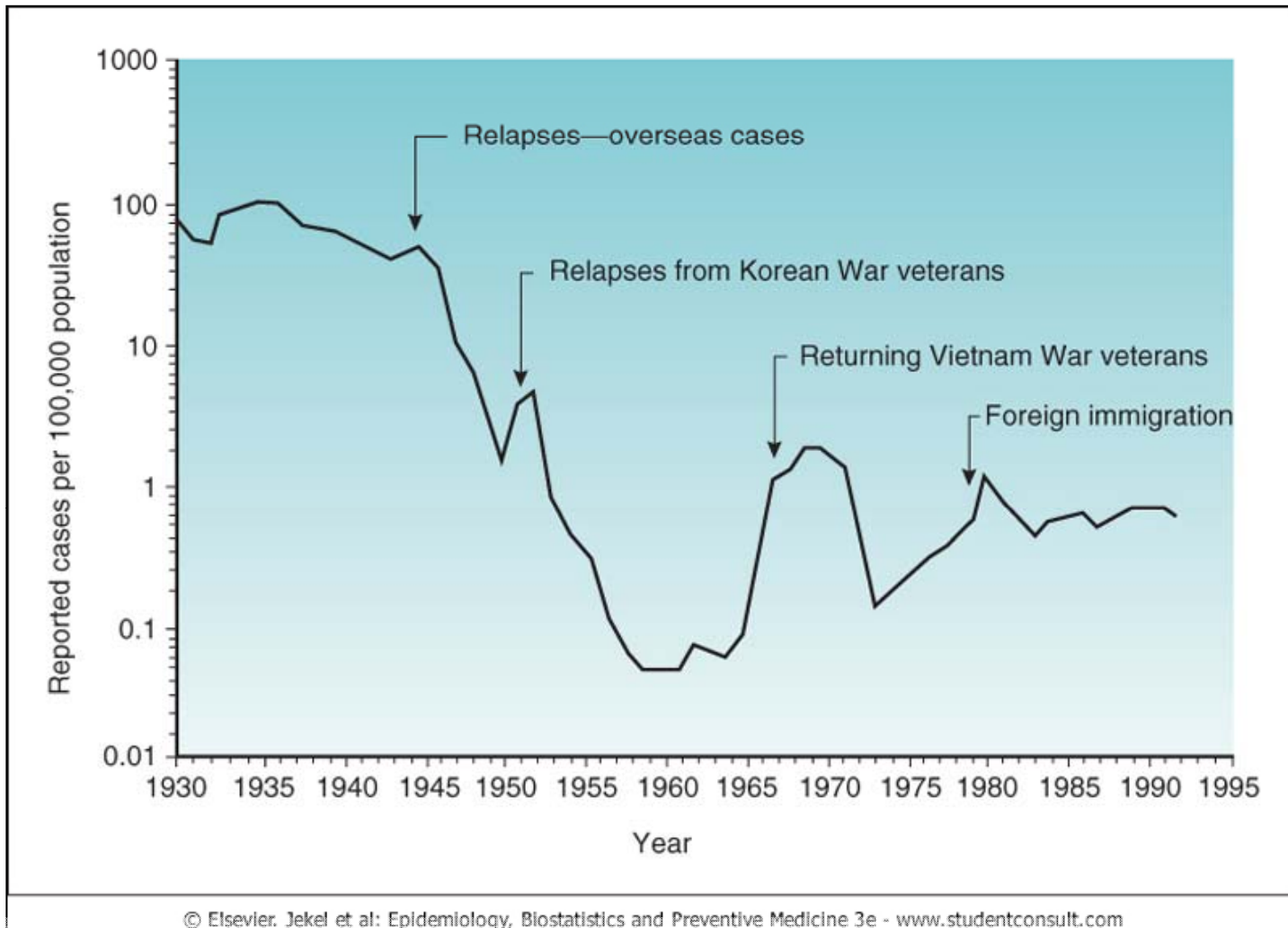
# Cross-Sectional Ecologic Studies

- Relate the frequency with which some **characteristic** (e.g., smoking, water fluoride) and some **outcome** of interest (e.g., lung cancer, dental caries in children ) occur in the **same geographic area** (e.g., a city or country).
- Useful for hypotheses generation, but they **cannot** be used to draw **causal** conclusions.
- Concerned citizens sometimes are unaware of these weaknesses (**ecologic fallacy**).



# Longitudinal Ecologic Studies

- Use ongoing surveillance or frequent cross-sectional studies to measure **trends** in disease rates over many years in a defined population.
- By comparing the **trends** in disease rates with other **changes** in the society (e.g., wars, immigration, or the introduction of a vaccine or antibiotics), epidemiologists attempt to determine **the impact of these changes on the disease rates.**



### (Longitudinal Ecologic Studies)

**Incidence rates of malaria in the US, by year of report, 1930-1992.**

**The peaks in malaria rates can be related to social events, such as wars and immigration.**

# Longitudinal Ecologic Studies

Important **causal** associations have been **suggested** by longitudinal ecologic studies.

(20 years after smoking rates in men ↑, the lung cancer rate in the male population ↑).

*National data on smoking and lung cancer rates*

The task of establishing a **causal relationship** was left to **cohort and case-control** studies.

# **Observational Designs for Generating or Testing Hypotheses**

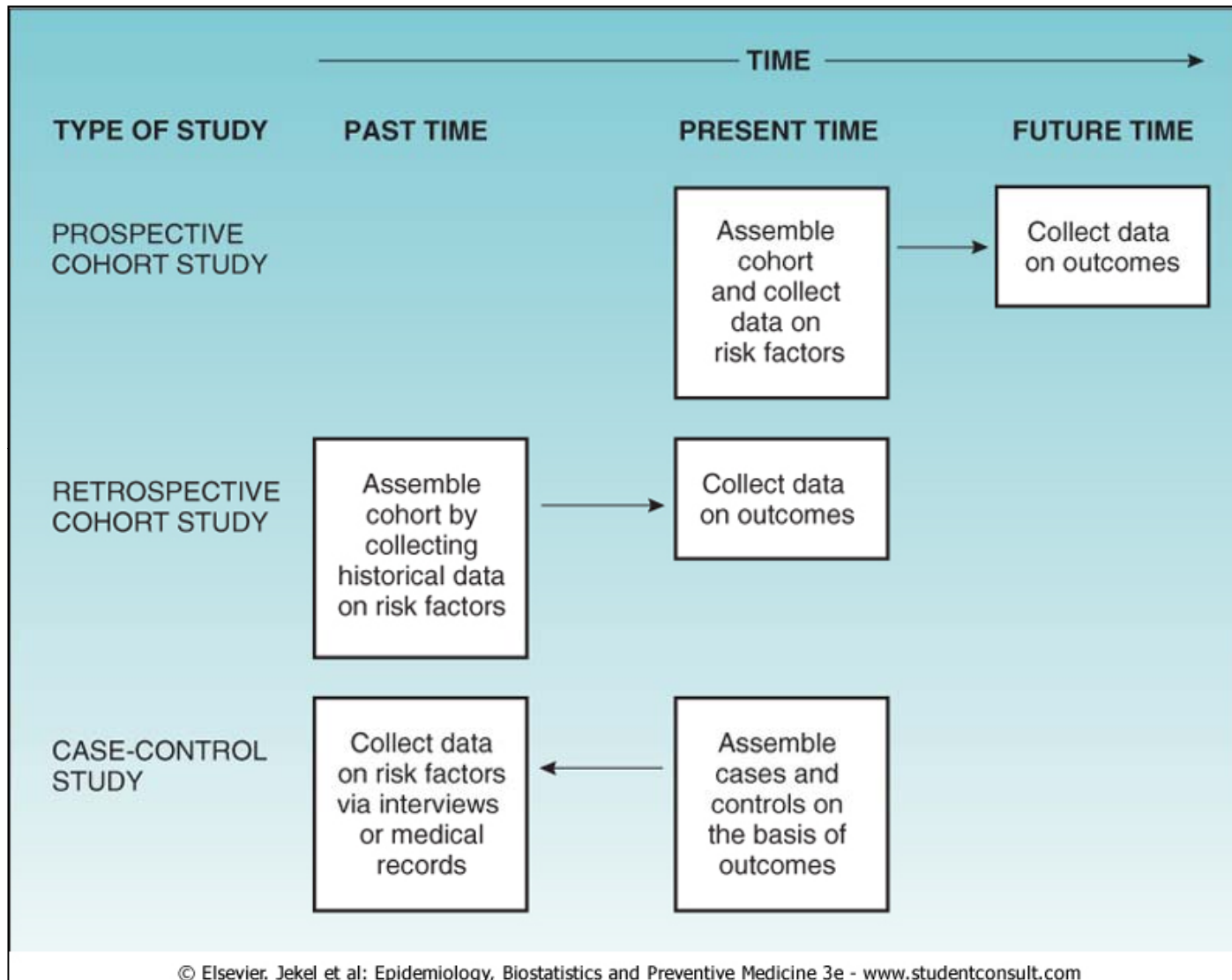
- 1. Cohort**
- 2. Case control**
- 3. Nested Case-Control**
- 4. Case- Crossover**

# Cohort Studies

- A **cohort** is a clearly identified group of people to be studied.
- In cohort studies, investigators begin by assembling **one or more cohorts**, either by **choosing persons specifically** because they were and were not exposed to one or more risk factors to be studied or by taking a **random sample** of a population.

# Cohort Studies

- The subjects of Cohort are followed over **time** to determine whether or not they **develop** the diseases of interest, and whether the **risk factors** that were measured at the **beginning** of the study predict the diseases that occur.
- **Cohort** are defined on the basis of **exposure** and are **followed for outcomes**. This is in contrast to **case-control** studies , in which groups are assembled on the basis of **outcome status** and are queried for **exposure** status.
- There are two general types of cohort study, the **prospective** type and the **retrospective** type.



**The relationship between the time of assembling study subjects and the time of data collection in a prospective cohort study, a retrospective cohort study, and a case-control study.**

# Prospective Cohort Studies

Assembles the study groups in the present time, collects baseline data on them, and continues to collect data for a period that can last many years.

## Advantages:

1. Investigator is able to control the data collection as the study progresses and can check the outcome events (e.g., diseases and death) carefully when they occur, ensuring that they are correctly classified.
2. The estimates of risk obtained from prospective cohort studies are true (absolute) risks for the groups studied.
3. Many different disease outcomes can be studied.



# Prospective Cohort Studies

## Disadvantages:

1. Only the risk factors defined and measured at the beginning of the study can be used.
2. High costs
3. Loss to follow-up
4. The long wait until their results are obtained.

The **Framingham Heart Study** (1950) and **the Nurses' Health Study** are Prospective Cohort Studies.

# Retrospective Cohort Studies

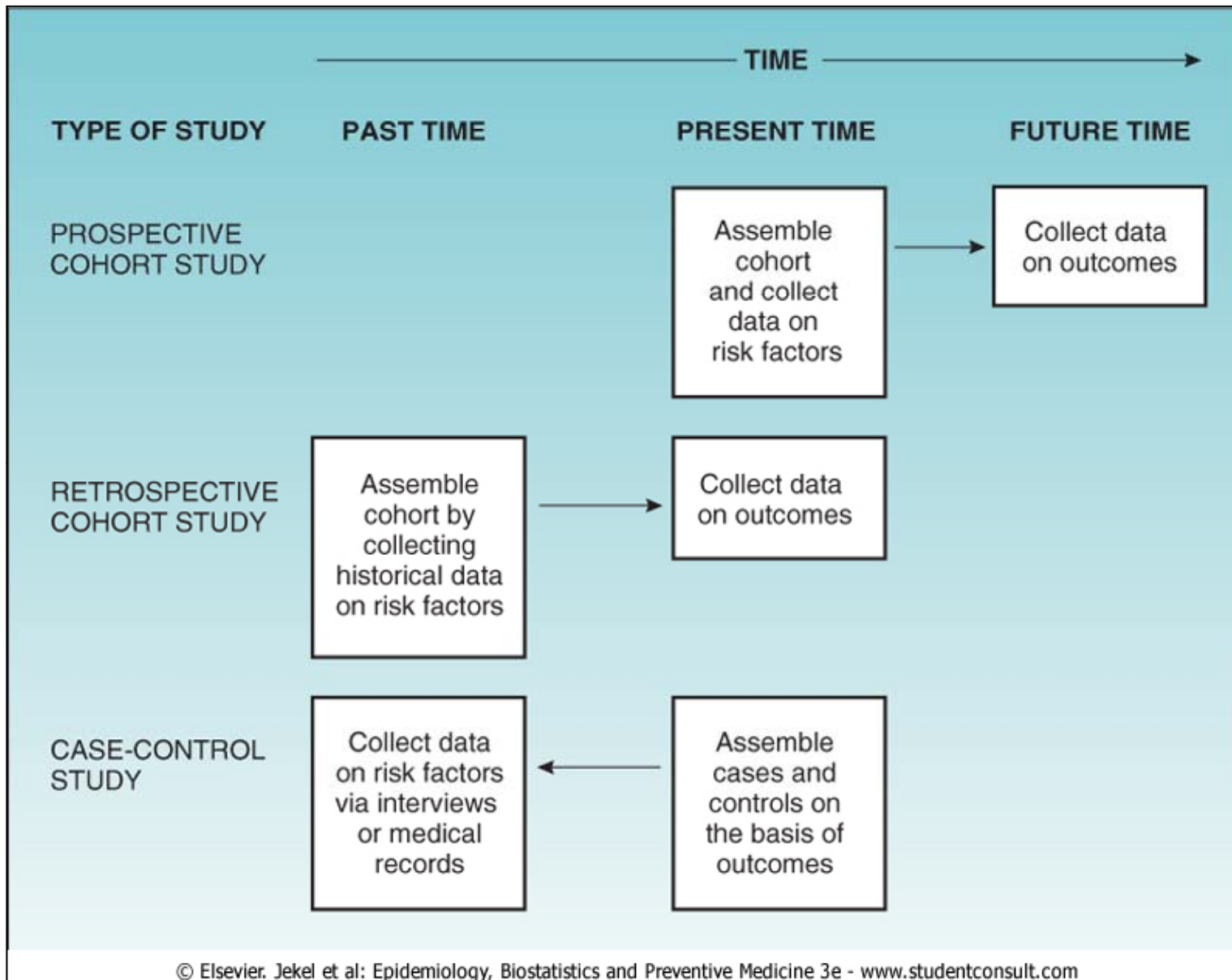
The investigator goes back into history to define a risk group (e.g., people exposed to the **Hiroshima atomic bomb in August 1945**) and follows the group members up to the present to see what outcomes (e.g., cancer and death) have occurred.

- Ability to calculate an absolute risk
- Lacks the ability to control the data collection process

A retrospective cohort study was done in **1962** to investigate the effects of **prenatal x-ray exposure**. They determined how many subjects from each group had gotten cancer during childhood or early adulthood (up to the time they did the study). They found that the individuals who had been exposed to x-rays in uterus had a 40% increase in the risk of childhood cancers (i.e., a risk ratio of 1.4).

# Case-Control Studies

- Case group and the control group are selected on the basis of the **outcome** (i.e., having the disease of interest versus not having the disease of interest) and **compares the groups** in terms of their frequency of **past exposure** to possible risk factors.



# Case-Control Studies

- The **actual risk** of the outcome **cannot** be determined from case-control studies because the underlying population is unknown.
- An **estimate** of the **relative risk** of the outcome, called the **odds ratio**, can be determined in case-control studies.

# Case-Control Studies

- The cases and controls are assembled, and they are questioned (or their relatives or medical records are consulted) regarding past exposure to risk factors.
- In past decades, case-control studies often were called **retrospective** studies for this reason.

# Case-Control vs. Cross Sectional

- The **time relationships** in a case-control study are similar to those in a **cross-sectional** study in that the investigator learns simultaneously about the current disease state and any risk factors that may have existed in the past.
- In terms of **assembling the subjects**, however, a case-control study differs from a cross-sectional study in that the sample for the case-control study is chosen specifically from groups with and without the disease of interest.

# Case-Control vs. Cohort

- **Case-control studies**

1. Quick
2. Inexpensive
3. Rare Diseases (e.g., prevalence of <1%)

- **Cohort studies**

1. Large sample size (With incidence 1 in 1000 persons per year, to find just 100 cases over 2 years, we should study 50,000 persons.)
2. Take a long time
3. Expensive
4. Loss to follow up

In case-control studies only **one outcome** (one disease) can be considered per study, however, **many risk factors** may be considered, and this makes case-control studies useful for **generating hypotheses concerning the causes of a disease.**



# Problems of Case-Control Studies

1. **Recall Bias**
  2. **The Correct Control Group** (The controls usually are **matched** individually to cases on the basis of **age, sex** and obtains from **the same diagnostic setting**. (e.g., pulmonary disease), so similar workup including a chest x-ray and spirometry have been done.)
  3. **Overmatching:** If cases and controls were matched on some characteristic that is potentially causal, that "cause" would be missed. (Lung cancer and match for smoking)
- Given the difficulties of selecting a control group with no bias, the investigator often assembles **two or more control groups, one of which is from the general population.**

# Nested Case-Control Studies

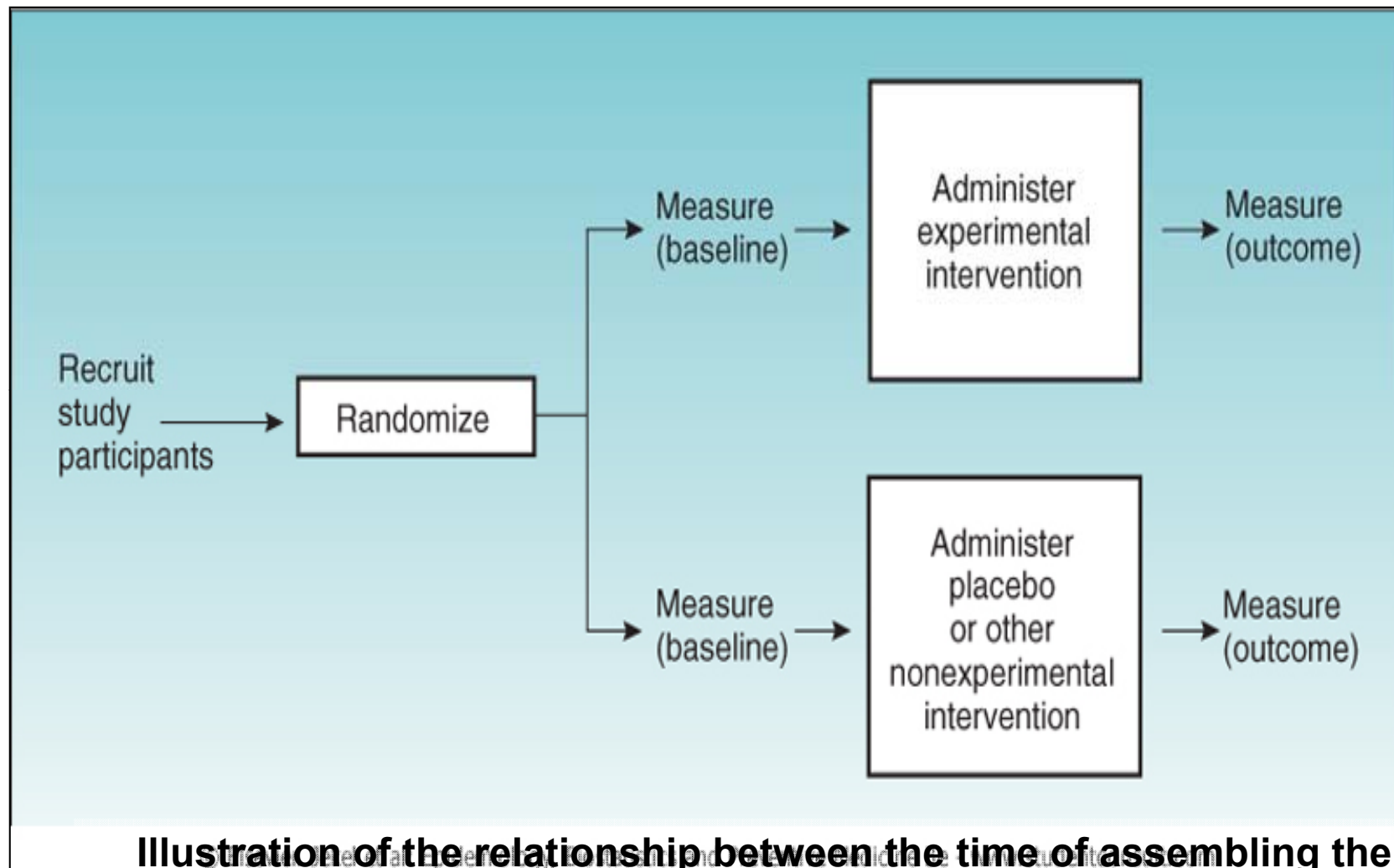
- Consists of a cohort study with a nested case-control study.
- In this new design, a **cohort of patients** (e.g; Diabetics or suspected meningitis ) is defined, and the baseline characteristics of the patients are obtained by interview, physical examination, and laboratory or imaging studies. The **patients are followed** to determine the outcome (e.g; MI or *nonbacterial* meningitis).
- Patients who develop the condition of interest become cases in a case-control study, and patients who do not develop the condition become eligible for the control group of a study. Next, the cases and a representative (or matched) sample of controls are studied, and data from the two groups are compared using **analytic methods appropriate for case-control studies**.

# Case-Cohort Studies

- A variant of the nested case-control study is known as the **case-cohort** study.
- In this approach, the study also begins with a cohort study, and the controls for the study are taken from the cohort study, but they are identified **before any cases develop (although some of them may become cases)**.
- The analysis here is more complex.

# Experimental Designs for Testing Hypotheses

1. Randomized controlled *clinical trials* (RCCTs)
  2. Randomized controlled *field trials* (RCFTs).
- Both designs follow the same series of steps (Fig) and have many of the same advantages and disadvantages.
  - The major difference between the two is that:  
**Clinical** trials are used to test **therapeutic** interventions in ***ill*** persons, whereas  
**Field** trials are done to test **preventive** interventions in ***well*** persons in the community.



**Illustration of the relationship between the time of assembling the study subjects and the time of data collection in an RCCT and an RCFT.**

# Randomized Controlled Clinical Trials

- In an RCCT, often referred to simply as randomized controlled trials (RCT), patients are enrolled in a study and **randomly** assigned to one of the following groups:
  - (1) the **intervention** group, which receives the experimental treatment, or
  - (2) the **control** group, which receives the nonexperimental treatment, consisting either of a placebo (inert substance) or of a standard method of treatment.

# Randomized Controlled Clinical Trials

- RCCTs are considered the "**gold standard**" for studying interventions because of their ability to **minimize bias** in the information obtained from the study subjects.
- Nevertheless, they do **not** entirely eliminate bias, and they pose some challenges and **ethical** dilemmas for investigators.

# Randomized Controlled Clinical Trials

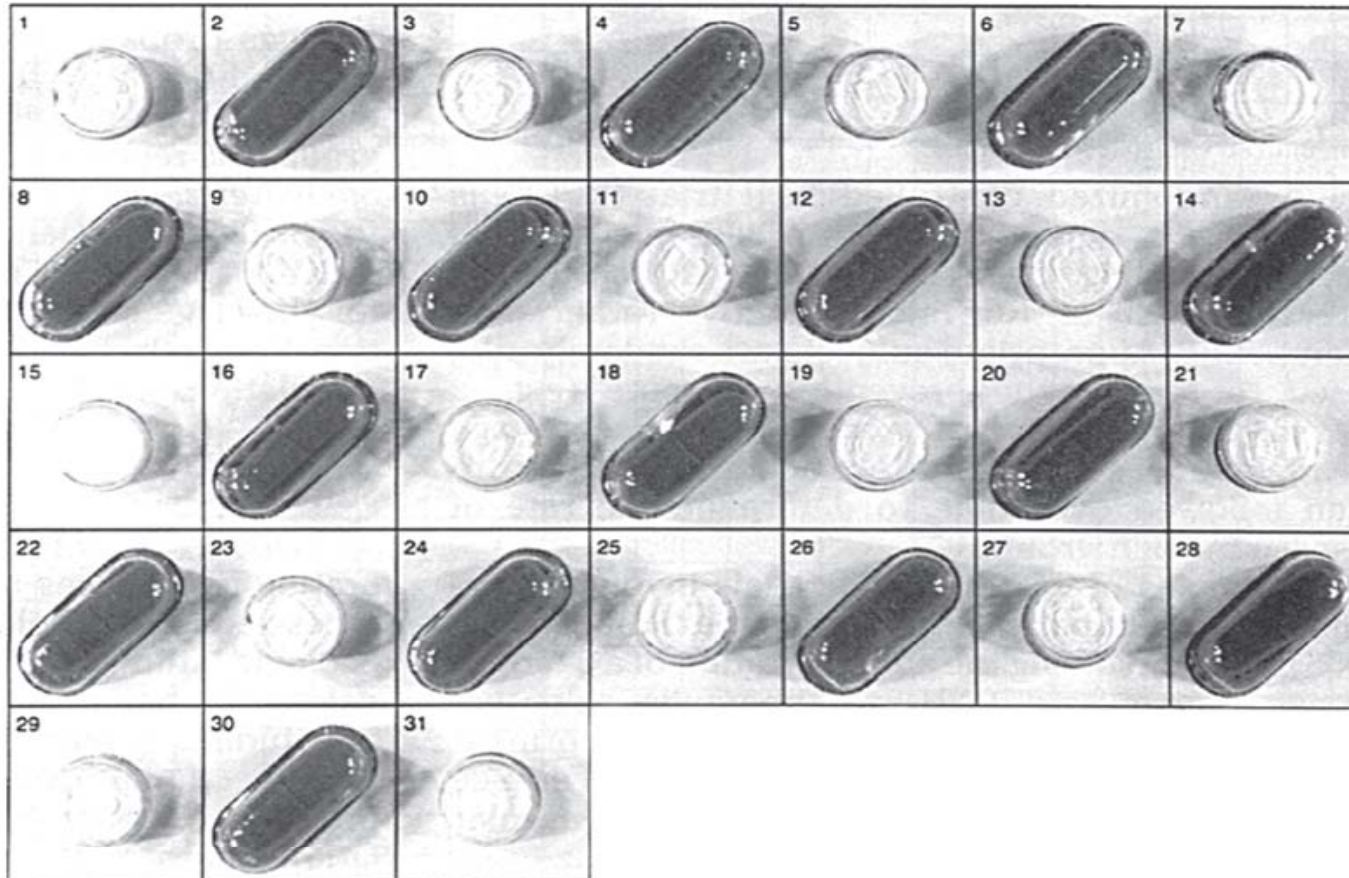
- To be enrolled in an RCCT, the **patients** must **agree** to participate **without knowing** whether they will be given the experimental or nonexperimental treatment; this would be a **single-blind study**.
- If possible, the **observers** who collect the data also are prevented from knowing which type of treatment each patient is given. When both of these protections are done, the trial is said to be a **double-blind study**.



Inquiries to: **PHYSICIANS HEALTH STUDY, HARVARD MEDICAL SCHOOL**

900 Commonwealth Avenue East, Boston, MA 02215

Directions: Please take 1 pill each day. Use the tablet or capsule with the current date.  
Store cards in a cool, dry place out of direct sunlight.



© Elsevier, Jekel et al: Epidemiology and Biostatistics, 3e - www.studentconsult.com

**22,000 physicians**

**Aspirin to reduce cardiovascular disease and beta carotene to prevent cancer.  
To have true blinding, the nonexperimental treatment must appear identical  
(e.g., in size, shape, color, taste) to the experimental treatment**

# Randomized Controlled Clinical Trials

- Blinding is impossible and unethical:
  1. **Surgical intervention**
  2. **Intervention were the best available**
  3. **Prenatal care**

# Problems of RCCT

- **Lost to follow-up** (for various reasons)
- **Therapy changes** (due to side effects)
- **Publication bias** (only positive results are publishing)

# Randomized Controlled Field Trials

- An RCFT is similar to an RCCT except that the intervention in an RCFT is **preventive** rather than therapeutic, and usually it is done in the **community**.
- Appropriate subjects are randomly allocated to receive the preventive measure (e.g., **a vaccine or an oral drug**) or to receive the placebo (e.g., **an injection of sterile saline or an inert pill**).
- They are followed over time to determine the rate of disease in each group.

# RCCT & RCFT

- **Disadvantages:**

1. The results may take a long time to obtain
2. Has to do with external validity

(which is the ability to generalize the findings to other groups in the population as opposed to **internal validity**, which concerns the validity of results for the persons in the study)

# Other Types of Study Design

- **Meta-analysis**

(summarize the information obtained in many single studies on one topic)

- **Decision analysis**

- **Cost-effectiveness analysis**

(summarize data and show how they can inform clinical or policy decisions)

# SUMMARY

- Research is the attempt to answer questions with valid data.
- Epidemiologists try to answer questions about:  
The state of health, disease, or risk factors;  
Develop hypotheses about :
  - 1) The causes of ill health
  - 2) The effectiveness of preventive and curative interventions and
  - 3) Test these hypotheses.

# SUMMARY

- Observational research designs suitable for generating hypotheses include cross-sectional surveys, cross-sectional ecologic studies, and longitudinal ecologic studies.
- A cross-sectional study collects data about a population at one point in time, whereas a longitudinal study is done over a period of time.
- Cross-sectional surveys are useful in determining the prevalence of risk factors and diseases in the population, but they are weak in determining the temporal relationship between variables.
- In ecologic studies, the rate of a disease and the frequency of exposure to a risk factor are obtained for an entire population, but the unit of study is the population and not individuals within it, so the exposure and the disease cannot be linked in individual persons.



# SUMMARY

- Observational research designs suitable for generating or testing hypotheses include **prospective** cohort studies, **retrospective** cohort studies, and **case-control** studies.
- For **cohort** studies, one group consists of persons exposed to risk factors, whereas another group consists of persons not exposed. The groups are studied to determine and compare their rates of disease.
- For **case-control** studies, the case group consists of persons who have a particular disease, and the control group consists of persons who do not have the disease, but are matched individually to the cases (e.g., in terms of age, sex, and type of medical workup). Each group is studied to determine the frequency of past exposure to possible risk factors.
- In case-control studies, the **relative odds** that a disease is linked with a particular risk factor (**odds ratio**) can be calculated.

# SUMMARY

- The use of a cohort study with a **nested case-control** design may enable some hypotheses to be tested quickly and cheaply.
- The experimental designs suitable for testing hypotheses are RCCTs and RCFTs.
- The major difference between RCCTs and RCFTs is that clinical trials generally are used to test therapeutic interventions, whereas field trials usually are done to test preventive interventions.
- A trial is called a double-blind study if neither the subjects who participate in it nor the observers who collect the data know which type of intervention each participant is given.



شیراز پاتوق  
ShirazPatoGh.Com

### تصویری از آخرین لحظات حیات پروفیسر حسینی

مسنورهای تپش قلب روشن است. شگفت اینکه در چنین حالتی در کمال حیرت پزشکان و متخصصین بیمارستان کانتونال دانشگاه زتو، پروفیسور حسینی در آخرین لحظات حیات به چیزی جز مطالعه و افزایش دانش خویش نمی‌اندیشد.

این تصویر منحصر به فرد را یکی از کارکنان خود بیمارستان به عنوان یک تصویر نکان‌دهنده و تاثیر گذار ثبت کرده است.

شام در کنار تخت استاد سرد شده است. ملاحظه دیگر نیازی به خوردن غذا نیست، پزشکان و مسولان بیمارستان دانشگاه به این نتیجه رسیده‌اند که معالجه روی قلب استاد دیگر اثری ندارد.

لذا آنژیوپکت تزریق چند دارو برای ادامه تپش قلب از رگ دست راست و آنژیوپکت تزریق مسکن درد از دست چپ ایشان را خارج و حتی ماسک تامین اکسیژن که دیگر ریه‌ها قادر به تامین آن نبود را برداشته‌اند و تنها