

# Cohort studies

Training Course in Sexual and  
Reproductive Health Research  
Geneva 2015

*Nguyen Thi My Huong, MD PhD*  
*WHO/RHR/SIS*



World Health Organization



Reproductive Health and Research



UNDP • UNFPA • WHO • World Bank  
Special Programme of Research, Development  
and Research Training in Human Reproduction

# *OUTLINE*

- Overview
- Definitions
- Study design
- Basic measures
- Advantages and disadvantages
- When to apply a cohort design
- Practical considerations



# Overview

## Two major categories of Epidemiological studies:

- **Observational studies:**

- Cohort studies
- Case-control studies
- Cross-sectional study
- *Have no control over exposures, simply observe what happens to groups of people.*
- *Examine associations between risk factors and outcomes*

## **Experimental studies**

- *Randomized controlled trials (RCT)*
- *Non-randomized trial*
- *Explore the association between interventions and outcomes.*



# Definitions

## ● Cohort:

- A group of individuals who have characteristics in common
- **Examples of cohorts:**
  - **Birth cohort:** all individuals in a certain geographic area born within a given period of time (usually a year).
  - **Marriage cohort:** All persons married within a given period of time
  - **Exposure cohort:** individuals assembled as a group based on some common exposure (e.g. radiation exposure during desert testing, smoking exposure...)



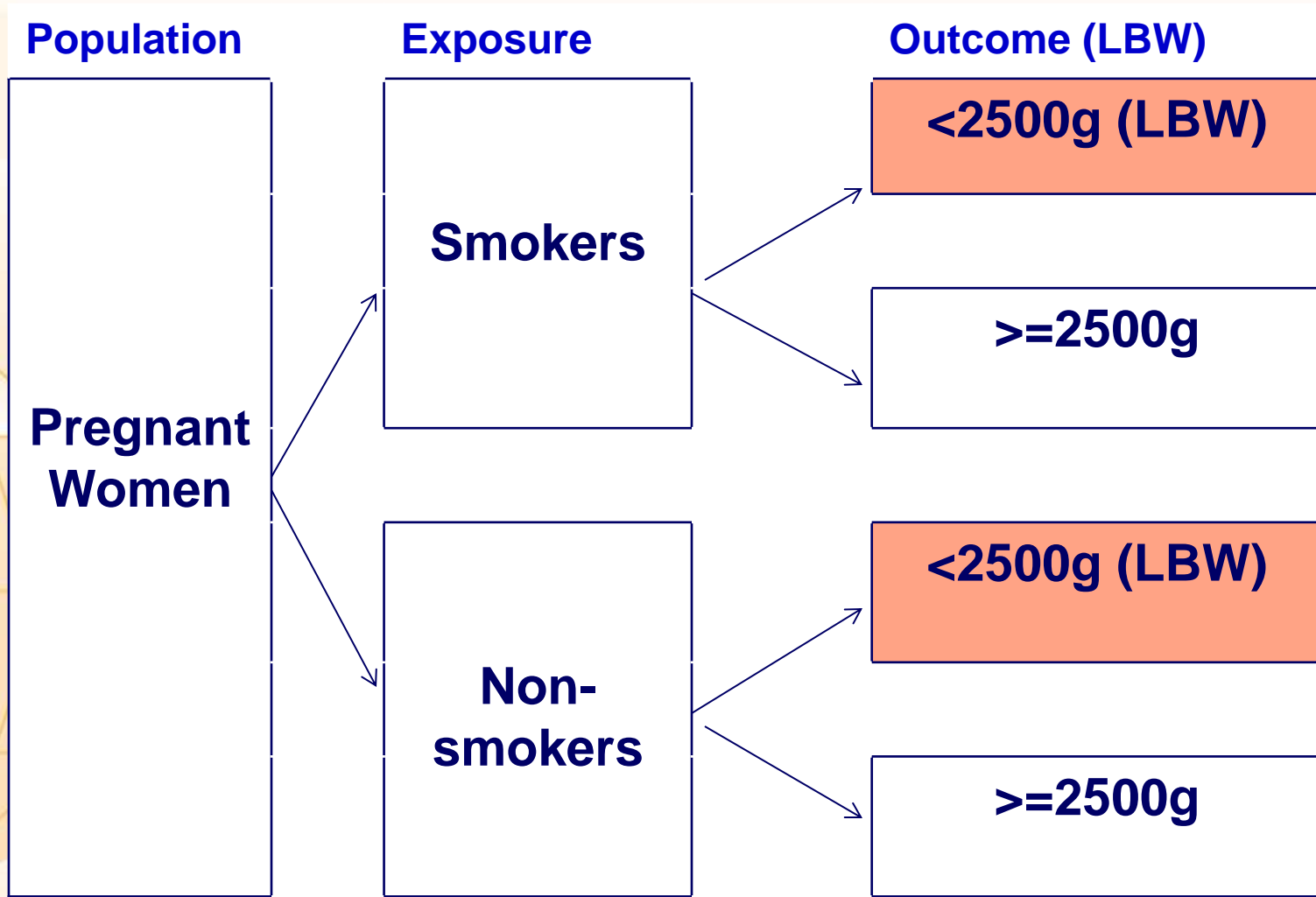
# Definitions

- **Cohort study:**

A study in which two or more groups of individuals those are free of disease and those differ according to the extent of exposure to a factor of interest, are followed over a period of time to see how their exposures affect their outcomes.



# Study design



# Type of cohort studies

- **Prospective cohort studies**
- **Retrospective cohort studies**
- **Classification is based on the temporal relationship between the initiation of the study (sample defined) and occurrence of the outcome, i.e. outcome before initiation (retrospective)**
- **Both start by identifying subjects based upon the presence or absence of the exposure of interest, without knowing the outcome at the time their exposure status is defined**



# Prospective cohort studies

- Sample defined prospectively during or before exposure and before outcome occurrence

## **Example:**

*(Ramchand R, Ialongo NS, Chilcoat HD. The Effect of Working for Pay on Adolescent Tobacco Use. Am J Public Health. 2007 Nov;97(11):2056-62.)*

- Cohort: High school students from Baltimore, Maryland
- Exposure: Working for pay
- Outcome: Initiation of tobacco use
- Results: Adolescents who work for pay have a higher risk of initiating tobacco use



World Health Organization



Reproductive Health and Research



UNDP • UNFPA • WHO • World Bank  
Special Programme of Research, Development  
and Research Training in Human Reproduction



# Prospective cohort studies

## **Example:**

*(Doll R, Hill AB. Mortality in Relation to Smoking: Ten Years' Observations of British Doctors. Br Med J. 1964 Jun 6;1(5396):1460-7.)*

- Cohort: British doctors responding to a survey in 1950
- Exposure: smoking
- Outcome: Lung cancer
- Periodic follow-up and review of death records
- Results: Smoking increased risk of lung cancer



World Health Organization



Reproductive Health and Research



UNDP • UNFPA • WHO • World Bank  
Special Programme of Research, Development  
and Research Training in Human Reproduction

# Prospective cohort studies

## **Example:**

*(Selikoff IJ, Hammond EC, Seidman H. Latency of asbestos disease among insulation workers in the United States and Canada. Cancer. 1980 Dec 15;46(12):2736-40.)*

- **Exposed:** 17,800 males in Asbestos Insulation Workers union in North America
- **Unexposed :** General population of males matched by age
- **Outcome:** Lung cancer
- **Results:** Positive association between asbestos and lung cancer



# Prospective cohort studies

## **Example:**

*(Nichol KL, Nordin JD, Nelson DB, Mullooly JP, Hak E. Effectiveness of Influenza Vaccine in the Community-Dwelling Elderly. New England Journal of Medicine. 2007 Oct 4;357(14) 1373-81.)*

- Exposed: Vaccinated elderly
- Unexposed: Unvaccinated community-dwelling elderly
- Outcome: Hospitalization for pneumonia or influenza
- Results: The elderly who were vaccinated have a reduced risk of hospitalization for pneumonia or influenza



# Retrospective cohort studies

- Both exposure and disease have occurred at the start of study.
- Data already collected for other purposes.
- The cohort is followed up retrospectively.
- It depends on the availability of previous study factor information.
- It is more feasible for studying a disease with a long latent period.
- The study period may be many years but the time to complete the study is only as long as it takes to collate and analyse the data.



# Retrospective cohort studies

## **Example:**

*(Olaf H Klungel SRH. Lipid-Lowering Drug Use and Cardiovascular Events after Myocardial Infarction. The Annals of pharmacotherapy 2002;36(5):751-7.)*

- Begin study in 2000 using data already collected via health plan.
- Cohort surviving myocardial infarction (MI) 1986-1996
- Exposed: Lipid lowering therapy use
- Outcome: Cardiovascular events during 6 months following MI



# Basic measures

- **Measures of disease occurrence:**
  - Cumulative Incidence
  - Incidence Rate (IR)
- **Measures of association between a factor and a disease:**
  - Relative Risk (RR)
  - Attributable Risk (AR)



# Basic measures

- ***Cumulative Incidence:***
  - Risk of developing disease
  - # new cases of disease/# persons at risk (during the same time period)



# Basic measures

## *Cumulative Incidence:*

- Risk of disease in exposed:  $a/a+b$
- Risk of disease in non-exposed:  $c/c+d$

	Disease	Non-disease	
Exposed	a	b	a + b
Non-exposed	c	d	c + d
	a + c	b + d	





# Basic measures

- ***Incidence Rate (IR)***

- Risk per unit of time
- # new cases of disease/Persons at risk\*Duration
- *Duration (Person-time)*: sum of time at risk for all individuals (time until the date of the event of interest or date of censoring, i.e. death, end of FU, drop out). e.g. 1 person FU for 2 years=2 person-year.
- *Persons “at risk”* who do not have the disease of interest and are capable of developing the disease.



# Basic measures

## Example:

(IR, Person-time calculation, a 9-year follow-up study)

Subject	Years of follow-up	Outcome
1	→ 2.1	Event
2	→ 4.8	Die
3	→ 3.2	Die
4	→ 9.0	End of FU
5	→ 7.2	Event

- *Person time:*  $2.1+4.8+3.2+9.0+7.2=26.3$  years
- *Incidence rate:*  $2 \text{ events}/26.3 \text{ person-years}=0.076/\text{year}$  (or 76/1000/year)



# Basic measures

- **Relative Risk (RR):**
  - Incidence of disease in exposed compared to the incidence of disease in unexposed
  - $RR = (a/a+b)/(c/c+d)$

	<b>Disease</b>	<b>Non-disease</b>	
Exposed	a	b	a + b
Non-exposed	c	d	c + d
	a + c	b + d	



# Basic measures

## – **Relative Risk (RR):**

- Determine the strength of the association between exposure and disease
- $RR=1$  (no association)
- $RR>1$  (exposure increases risk for disease, e.g.  $RR=2.0$  can be interpreted as two fold increase in risk)
- $RR<1$  (exposure decreases risk for disease, e.g.  $RR=0.7$  can be interpreted as 30% decrease in risk)



# Basic measures

## Example:

*(Tuberculosis treatment and breast cancer study)*

- Exposed: women were treated with air collapse therapy and exposed to numerous fluoroscopic examinations (radiation)
- Unexposed: women who received other treatment.
- Outcome: A total of 47036 woman-years of follow-up were accumulated during which 56 breast cancer cases occurred



World Health Organization



Reproductive Health and Research



UNDP • UNFPA • WHO • World Bank  
Special Programme of Research, Development  
and Research Training in Human Reproduction

# Basic measures

## Example:

*(Tuberculosis treatment and breast cancer study)*

	Breast Cancer	Non-disease	Total	Women-years of FU
Exposed	41	1006	1047	28,011
Non-exposed	15	702	717	19,025
	56	1708	1764	47,036

- $IR_{\text{exposed}} = 41/28011 = 1.5/1000$  woman-years
- $IR_{\text{non-exposed}} = 15/19025 = 0.8/1000$  woman-years
- $RR = IR_{\text{exposed}}/IR_{\text{non-exposed}} = 1.9$
- Results: Women exposed to fluoroscopies had 1.9 times the risk of breast cancer compared to unexposed women.



# Basic measures

- ***Attributable Risk (AR):***

- The excess risk of disease observed among exposed subjects.
- $AR = IR_{\text{exposed}} - IR_{\text{non-exposed}}$

## **Example:**

*(Tuberculosis treatment and breast cancer study)*

- $IR_{\text{exposed}} = 1.5/1000$  woman-years
- $IR_{\text{non-exposed}} = 0.8/1000$  woman-years
- $AR = IR_{\text{exposed}} - IR_{\text{non-exposed}} = 1.5 - 0.8 = 0.7/1000$  w/y
- Excess IR of breast cancer among women exposed to fluoroscopies was 0.7/1000 woman-year



# Advantages

- Gold standard for studying the association between risk factor and outcome
- Useful for looking at multiple exposures and their interactions
- Can evaluate multiple outcomes /diseases
- Clear time sequence (temporal relationship between exposure and outcome) strengthens the inference about cause





# Advantages

- Less bias due to prospective evaluation of exposures
- Efficient for rare exposures
- The best or only ethical way, sometimes, to do the study (situations where randomization is not possible)



# Disadvantages

- Time consuming
- The problem of attrition: loss of subjects (e.g. migration or death from other causes)
- Unexpected changes over time:
  - Changes to the environment can influence the association of disease and possible cause
  - Changes in diagnostic criteria and methods
  - Changes of staff
- Financial problems: lack of funding and the high costs of record keeping



# When to apply a cohort design

- In many cases, *cohort studies are preferred to RCT because they do not require strict random assignment of subjects, which is unethical or improbable.*
- Sometimes they are the only methods available. (e.g. testing the effect of smoking on health, random assignment would be infeasible and unethical. A *reasonable alternative* would be a cohort study with *two groups smokers and non-smokers and follows them forward through time to see what health problems they develop.*



# Practical considerations

- **Selection of comparable groups:**
  - Select a comparison (unexposed) group as similar as possible to the exposed group with respect to all factors except the exposure
- **Comparable ascertainment of the outcome in both groups:**
  - Blind the investigator conducting follow-up and confirming the outcome



# Practical considerations

- **Minimize “lost to follow-up”**
  - Exclude those likely to become “lost” (e.g. Planning to move, unwilling to return)
  - Obtain complete tracking information (address, phone number of subjects as well as of close friends and relative)
  - Maintain periodic contact (reminders, updates)



# References

- Doll R, Hill AB. Mortality in Relation to Smoking: Ten Years' Observations of British Doctors. Br Med J. 1964 Jun 6;1(5396):1460-7.
- Nichol KL, Nordin JD, Nelson DB, Mullooly JP, Hak E. Effectiveness of Influenza Vaccine in the Community-Dwelling Elderly. New England Journal of Medicine. 2007 Oct 4;357(14):1373-81.  
<http://dx.doi.org/10.1056/NEJMoa070844>
- Olaf H Klungel SRH. Lipid-Lowering Drug Use and Cardiovascular Events after Myocardial Infarction. The Annals of pharmacotherapy. 2002;36(5):751-7.  
<http://dx.doi.org/10.1345/aph.1A308>
- Ramchand R, Ialongo NS, Chilcoat HD. The Effect of Working for Pay on Adolescent Tobacco Use. Am J Public Health. 2007 Nov;97(11):2056-62.  
<http://dx.doi.org/10.2105/AJPH.2006.094045>
- Selikoff IJ, Hammond EC, Seidman H. Latency of asbestos disease among insulation workers in the United States and Canada. Cancer. 1980 Dec 15;46(12):2736-40.





***THANK YOU  
VERY MUCH***



World Health Organization



Reproductive Health and Research



UNDP • UNFPA • WHO • World Bank  
Special Programme of Research, Development  
and Research Training in Human Reproduction