Strategies for data analysis: case-control studies

Postgraduate Research Training in Reproductive Health

Dr Calvin Tiyou Kenmeni
Department of Obstetrics and Gynaecology
Faculty of Medicine and Biomedical sciences (FMBS)
University of Yaounde I – Cameroon
tiyoukc@yahoo.fr

Yaounde, 14 November 2007

Two main types of epidemiologic studies

- **Observational** : the epidemiologist observes the association between exposure and outcome (e.g. passive smoking and breast cancer)

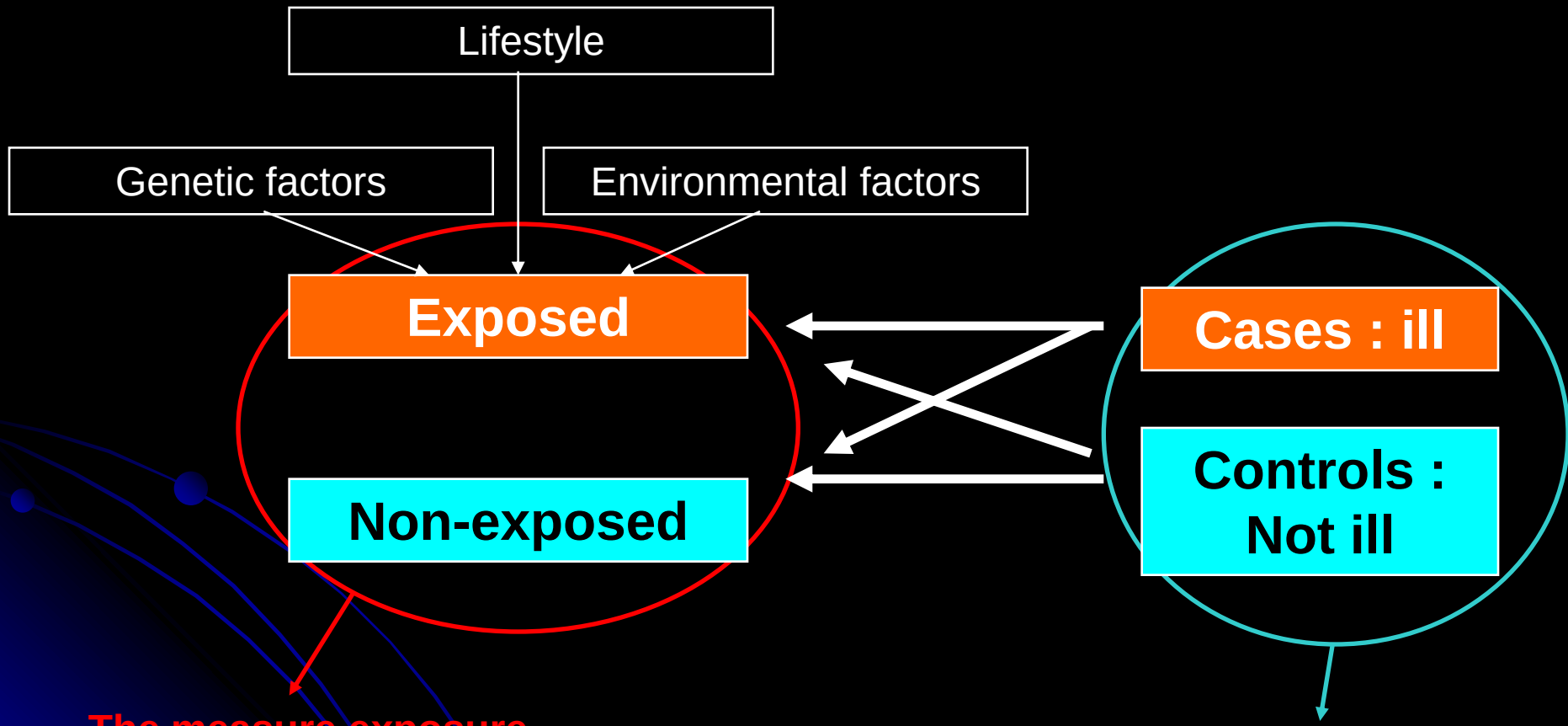
- **Experimental** : the epidemiologist performs an experiment, he/she controls the conditions under which the study is conducted (he/she is able to assign subjects to a treatment or comparison group and then follow them up to see if there are differences in the occurrence of disease between the two groups; e.g. calcium supplementation and pre-eclampsia)

Two types of observational studies

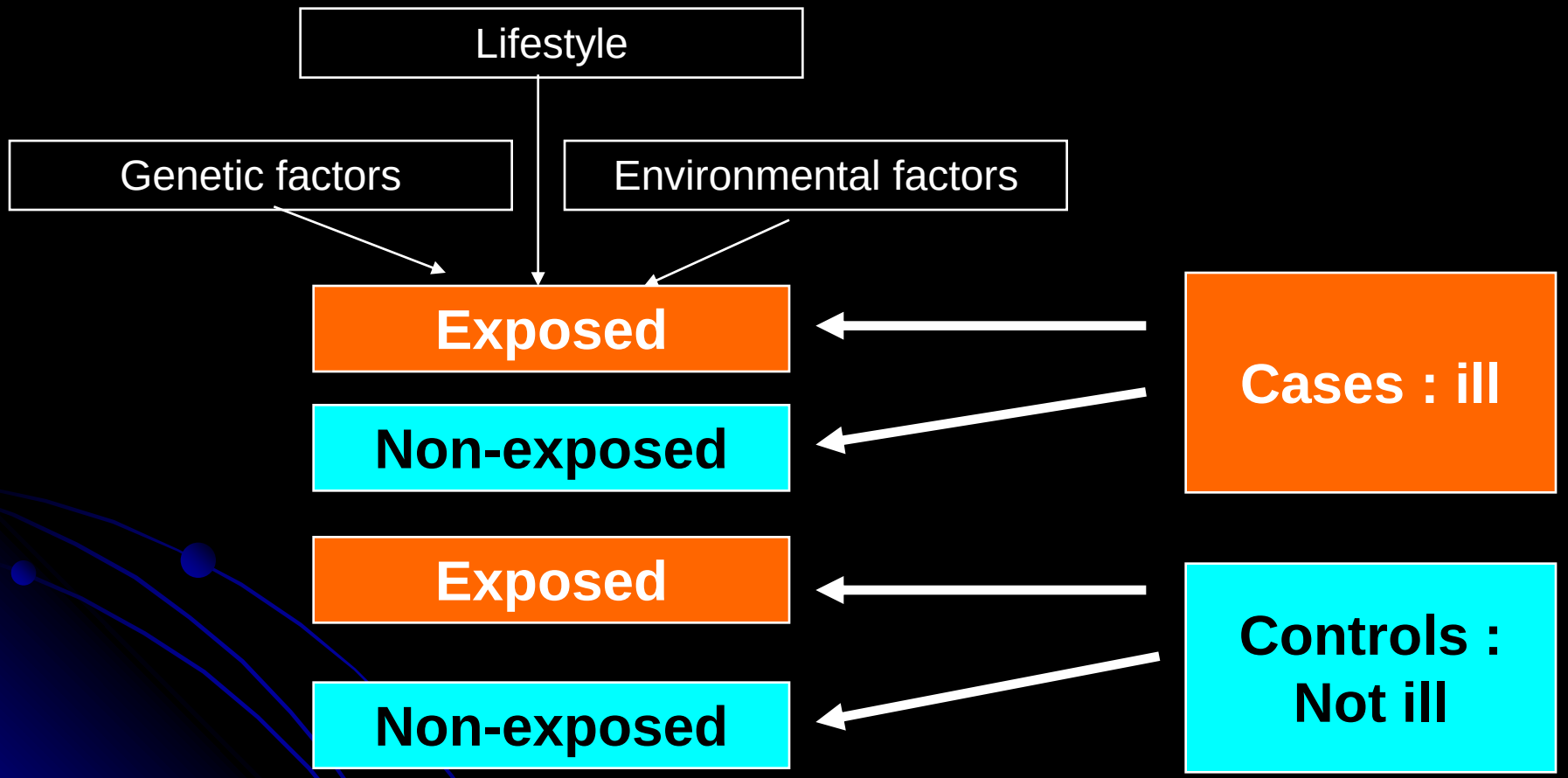
- **Cohort study** : one begins with a group of persons exposed to a factor of interest and a group of persons not exposed. The persons are then followed for the development of the disease of interest.

- **Case-control study** : one assemble a group of persons with a disease (cases) and a group of persons with no disease (controls). The history of past exposure to the factor of interest is then compared between the cases and controls.

Design of case- control studies: retrospective



Design of case- control studies: retrospective



Did they were exposed or not ?

Prevalence of disease is fixed by design

	Cases	Controls
Exposed	a	b
Non-exposed	c	d

Fixed margins

a+c and b+d

Prevalence of exposure in cases and controls and odds of being exposed

In case-control studies we can calculate :

- Prevalence of exposure in cases and in controls

  a/(a+c) and b/(b+d)

- The odds ratio to measure association between disease and exposure:

  The odds of being exposed for a case is a/c

  The odds of being exposed for a control is b/d

- The odds ratio of exposed vs non-exposed is

  OR = (a/c)/(b/d) = (a x d)/(b x c)

We can not calculate the relative risk (RR)

Interpretation of the odds ratio

- If exposure and disease are not associated, OR=1
- If exposure and disease are positively associated, OR>1
- If exposure and disease are negatively associated, OR<1

The OR is a good estimation of the RR if the disease is rare **(prevalence < 10%)**

Strategy for data analysis for case-control studies

- Describe study profile : number of cases and controls, identified and analyzed
- Baseline characteristics of cases and controls
- Crude ORs for different categories of use and risk factors
- ORs for different categories of use and risk factors, adjusting for confounders

Example
Oral contraceptives and breast cancer

Lancet 1985; 326:970-972

Study aim was to investigate relation between use of oral contraceptives (OCs) by young women and their risk of breast cancer

Cases: women 20-44 years at initial diagnosis of breast cancer, between Dec 1 1980 and Dec 31 1982, resident in 8 regions of the US, identified from population-based cancer registries.

Controls: women 20-44 years selected during same 25 months as the cases were diagnosed, residents of the 8 regions, selected randomly by telephone calls to households.

Fundamental conditions for the validity of this case-control design

Cases and controls :

- resident in 8 regions of the US, 20-44 years.
- are sampled independently of their use of oral contraceptives (OCs)

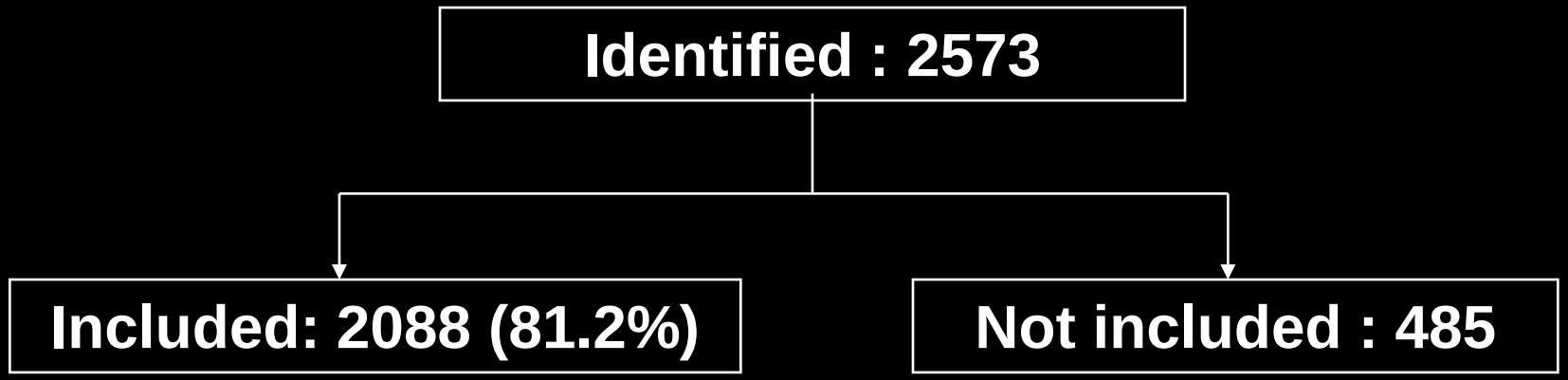
Solution :

- All incident cases over a given time period (Dec 1 1980 and Dec 31 1982)
- Controls are a random sample of population

Strategy for data analysis for case-control studies

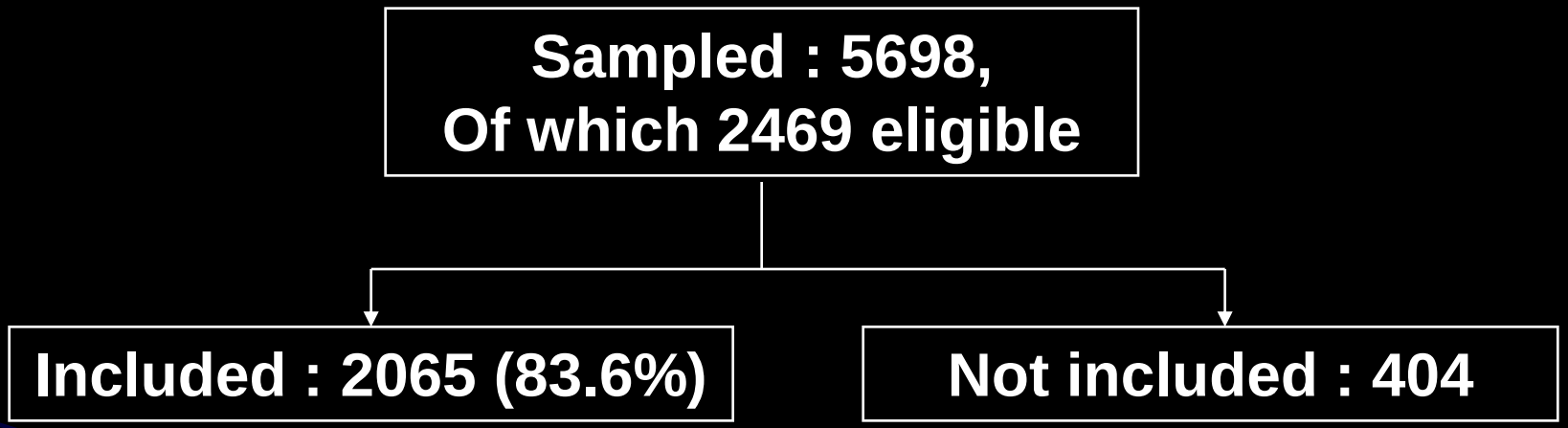
- Describe study profile : number of cases and controls, identified and analyzed
- Baseline characteristics of cases and controls
- Crude ORs for different categories of use and risk factors
- ORs for different categories of use and risk factors, adjusting for confounders

Oral contraceptives and breast cancer
Study profile : cases



6.4% refused to participate
3.7% died or were too ill
8.7% miscellaneous reasons

Oral contraceptives and breast cancer
Study profile : Controls



11.2% refused to participate
2.2% moved out
3.0% miscellaneous reasons

Strategy for data analysis for case-control studies

- Describe study profile : number of cases and controls, identified and analyzed
- Baseline characteristics of cases and controls
- Crude ORs for different categories of use and risk factors
- ORs for different categories of use and risk factors, adjusting for confounders

Oral contraceptives and breast cancer
Baseline characteristics of cases and controls

	Cases (%)	Controls (%)
Age		
20-24	0.7	5.1
25-29	6.0	8.2
30-34	18.3	20.8
35-39	33.5	28.6
40-44	41.4	37.3
Family history of breast cancer		
Yes	29.1	18.7
No	42.3	51.4
Unknown	28.6	29.9

Oral contraceptives and breast cancer
Baseline characteristics of cases and controls

	Cases (%)	Controls (%)
Age at first term pregnancy		
Nulliparous	18.2	18.4
< 20	19.1	22.1
20-22	23.1	24.7
23-26	22.3	21.6
27-29	9.2	7.6
➢29	7.0	4.0
Parous Unknown age	1.1	1.6
Benign breast disease surgery		
Yes	4.5	2.3
No	87.2	91.7
Unknown	8.3	6.0

Strategy for data analysis for case-control studies

- Describe study profile : number of cases and controls, identified and analyzed
- Baseline characteristics of cases and controls
- Crude ORs for different categories of use and risk factors
- ORs for different categories of use and risk factors, adjusting for confounders

Oral contraceptives and breast cancer
Results : crude ORs

	Cases	Controls
Exposed	1701	1662
Non-exposed	387	403
All	2088	2065

OR=1.07

Stratification and confounding variables ?

Oral contraceptives and breast cancer Results : crude ORs

Age at diagnosis or selection	Cases (%)		Controls (%)		OR
	N	% ever users	N	% ever users	
20-24	15	100.0	106	69.8	-
25-29	126	86.5	169	87.6	0.91
30-34	382	89.3	429	88.1	1.13
35-39	700	886.4	590	85.1	1.11
40-44	865	73.0	771	72.6	1.02
Total	2088	81.5	2065	80.5	1.07

Strategy for data analysis for case-control studies

- Describe study profile : number of cases and controls, identified and analyzed
- Baseline characteristics of cases and controls
- Crude ORs for different categories of use and risk factors
- ORs for different categories of use and risk factors, adjusting for confounders

Oral contraceptives and breast cancer Results : adjusting ORs

We need to adjust for factors associated with the risk of breast cancer or with the likelihood of diagnosis :

- Age at diagnosis or selection
- Family history of breast cancer
- Age at first term pregnancy
- History of surgery for benign breast disease
- Frequency of breast examination

Techniques to adjust ORs :

- Logistic regression
- Mantel-Haenszel

Oral contraceptives and breast cancer Results : adjusting ORs

Age at diagnosis or selection	Cases (%)		Controls (%)		OR (95% CI) (adjusted)
	N	% ever users	N	% ever users	
20-24	15	100.0	106	69.8	-
25-29	126	86.5	169	87.6	1.0 (0.5-2.1)
30-34	382	89.3	429	88.1	1.2 (0.7-1.8)
35-39	700	886.4	590	85.1	1.1 (0.8-1.6)
40-44	865	73.0	771	72.6	1.1 (0.9-1.4)
Total	2088	81.5	2065	80.5	

Oral contraceptives and breast cancer Conclusions

There was no significant increase or decrease in the risk of breast cancer for OC users according to

- Age at diagnosis
- Age at first use
- Duration of use
- Use before first term pregnancy

Use of OCs by young women in the US has no effect on the risk of breast cancer before 45 years of age.

Advantages of Case-Control Studies

- Less expensive …
- Require smaller sample sizes …
- Shorter duration …than prospective study
- Study multiple risk factors for 1 disease
- Easily reproduced in different populations by different investigators

Disadvantages of Case-Control Studies

- Information about exposure is often obtained after the diagnosis is done
  - Example: diet, physical activity
- Dependent on the subject's memory, which may be affected by the disease
- Population of origin for cases is difficult to define precisely
  - Difficult to identify appropriate control group
- Does not provide estimate of risks and attributable risk

Thank you for lending me your ears

Special acknowledgements to

Gilda Piaggio, *World Health Organization*
Hans Wolff, *Hôpitaux Universitaires de Genève (HUG)*