

환자-대조군 연구

**Case-Control Studies**

한림대학교 의과대학 김 동 현

# 학습 목표

1. 환자-대조군 연구의 정의, 방법과 특징을 설명한다.
2. 지역사회 기반과 병원기반 환자-대조군 연구의 장단점을 이해한다.
3. 환자-대조군연구에서 대조군 선정기준을 이해한다.
4. 환자-대조군 연구에서 폭로요인과 질병과의 관련성의 크기를 계량화하는 지표의 산출과정을 이해한다.

# 대표적인 연구 사례

---

- 폐암과 흡연 (Doll et al, Wynder et al, ... 1950's)
- Vaginal cancer 와 DES (diethylstilbestrol), (Herbst et al., 1971)
- 유방암과 첫 출산연령 (MacMahon et al., 1970)
- Endometrial cancer 와 Estrogen use (Ziel et al., 1975)

# 기본 개념 및 형태

---

- 질병이 있고 없음에 따라 연구대상자를 선정하고, 이들에게서 과거 폭로력을 조사, 비교하여, 해당 폭로요인과 질병과의 관련성을 평가한다.
- 만성질환과 상대적으로 발생빈도가 드문 질환의 연구에 주로 적용
- 연구의 흐름



	D	D	
E	a	b	a+b
E	c	d	c+d

# 연구의 설계 (1)

---

## □ 연구 주제의 선정

- 연구변수 결정 – 결과변수, 폭로변수, 기타 관련요인
- 적절한 연구 대상집단의 선정

## □ 환자군의 선정

- 질병에 대한 정의 – 객관적 기준, 재현성 있는 기준
- 신규환자 대 유병환자 ( **Incident vs prevalent cases** )
- 입적, 제외 기준 ( **Eligibility and exclusion criteria** )
  - 가능한 한 균질한 환자집단 (homogenous case group)
  - 환자, 대조군에 동일한 기준의 적용
  - 분명한 발병원인이 알려진 경우는 제외

# 연구의 설계 (2)

---

- 환자군의 종류

- 1) 병원환자

- 대상 선정이 용이하다
- 선택편견의 작용, 환자의 분포가 제한적

- 2) 질병 등록사업 또는 지역조사

- **full spectrum of diseases**
- 선택편견이 없다
- 연구수행이 어려움

# 환자-대조군 연구 - 연구의 설계 (3)

---

## □ 대조군의 선정

- 환자군을 생산해내는 모집단 (pool) 으로서의 대조군
- 연구 대상 질병이 없고, 그 질병에 이환될 가능성이 있어야
- 대표성(representativeness) 보다는 비교성(comparability)이 중요
- 환자군과 동일한 입적/제외 기준이 적용
- 대조군이 되기 위한 필요조건
  - 1) 연구중인 질병이 없어야 한다
  - 2) 연구가설상의 폭로요인에 노출가능하여야 한다

# 환자-대조군 연구 - 연구의 설계 (4)

---

- 대조군의 종류

- 1) 병원 대조군

- 환자와 동일한 병원에 입원한 타 질환 환자에서 선정
    - 환자군과 유사한 referral pattern 을 갖는 질환에서 선정
    - 주 폭로요인과 관련성이 알려져 있지 않은 질환에서 선정

- 장점 ; 대상자 확보가 용이

- 과거 폭로력에 대한 기억에 있어 환자군과 유사  
환자군과 같은 경로로 병원을 찾게 된다  
협조적이다.

- 단점 ; 폭로요인의 분포가 일반인구와 다를 수 있다

- 대조군이 갖고 있는 질환이 주 폭로요인과 관련 ?

# 환자-대조군 연구 - 연구의 설계 (5)

---

## 2) 지역 대조군

- random digit dialing, population census, voting list, ... .
- 환자가 나온 지역을 확정할 수 있어야
- 폭로유무에 따라 연구대상으로서의 선정에 비뚤림이 작용하지 않아야
- 비교성 ; 환자를 배출한 집단과 동일한 집단에서 선정
- 장점 ; 지역 환자군과는 비교성이 아주 좋다 (**high comparability**)  
선택편견이 배제
- 단점 ; 연구에 대한 태도가 상이 (**different mental setting**)  
응답율이 낮다  
비용이 많이 든다

## 3) 특수 집단 ; 이웃, 친지, 친구 등등

# 자료 수집 및 정리 (1)

---

- 폭로력과 기타 관련요인에 대한 조사
  - 환자군과 대조군에 대해 동일한 조사기법의 적용
    - 설문조사 : 면접 / 전화조사
    - 의무기록의 이용
  - 질병이 발생하기 전의 폭로경험에 대한 조사
    - 질병의 발병기전에 대한 생물학적 이해가 필요
    - 잠복기에 대한 고려
  - 인구학적 변수, 질병과거력, 약물복용력, 직업력, 가족력, 흡연/음주, 운동, 식이, 임신/출산력 ... .
  - 교란요인에 대한 정보가 없다면, 그에 대한 통제가 불가능

# 자료 수집 및 정리 (2)

---

- 분류오류에 의한 비뚤림 ( **misclassification bias** )
    - 정확도의 정도가 환자군과 대조군이 동일해야
    - 차별적 대 비차별적 ( **differential vs non-differential bias** )
    - 정보 비뚤림 ( **Information bias** ), 회상 비뚤림 ( **recall bias** )
  - 폭로요인에 관한 정보의 신뢰도와 타당도 검토
    - 식이요인
- 자료의 정리 ( **data cleaning** )
- 자료의 입력 – 데이터베이스 구축
  - 오류 확인 ( **error checking** )
  - **Missing information** 에 대한 처리
  - 내적 일치도의 확인 ( **internal consistency** )
  - 단변수 빈도표의 작성

# 자료 분석 (1)

---

- 관련도 지표 (Effect estimate) 산출
  - Odds ratio 와 95% 신뢰구간의 계산

	D	D	
E	a	b	a+b
E	c	d	c+d

- Odds ratio = 
$$\frac{\text{exposure odds among the diseased (}\Rightarrow a : c \text{)}}{\text{exposure odds among the undiseased (}\Rightarrow b : d \text{)}} = ad / bc \text{ ( null value 1 )}$$

# 자료 분석 (2)

---

- Odds ratio 와 95% 신뢰구간의 계산
  - $OR > 1$  ; risk factor,  $OR < 1$  ; protective factor
  - 95% 신뢰구간 (confidence interval, CI)
    - $\text{Log}(OR) \pm 1.96 \text{ SE of } [\text{log}(OR)]$
- 폭로수준이 여럿일 경우 ( Multiple level of exposure, 2 x R table)
  - 경향 분석 ( Test for trend )
- 층화 분석 (Stratified analysis)
- 다변량 로짓 분석 (Multivariable logistic regression)

# 결과의 해석

---

- 통계적 설명보다는 생물학적 해석이 보다 중요
- 연구 결과의 해석은 종합적이어야 한다.
  - 연구설계의 평가
  - 여러 비뚤림에 의한 영향 평가
  - 생물학적 가설
  - 다른 연구결과와의 비교
  - 연구의 제한점과 추가 연구가 필요한 부분

# 환자-대조군 연구의 장단점

---

## □ 장점

- 신속하고, 상대적으로 저렴하게 연구 수행
- 잠복기가 긴 질환의 경우 적절
- 발생이 드문 질환의 경우 바람직
- 여러 폭로요인에 대한 연구를 동시에 수행

## □ 단점

- 폭로요인에 대한 노출이 희귀할 경우 연구수행이 어렵다
- 폭로요인-질병간의 시간적 선후관계를 분명히 하기가 어렵다
- 선택편견과 정보편견이 잘 일어난다
- 질병 발생률을 직접 계산할 수 없다