

코호트 연구의 역사적 변천

고려대학교 의과대학 예방의학교실
김순덕

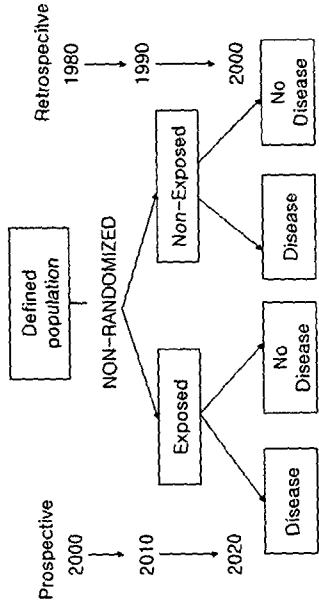
1. 코호트 설계
2. 설계의 역사적 변천
3. 분석방법의 역사적 변천

1. 코호트 설계

- 시간의 경과를 통한
- 연구시작부터 연구종료의 기간 동안 참여자 추적
- 결과 측정들을 관찰
: 사망, 질병발생, 생물학적 지표의 변화, 건강상태

- 전사(warriors)를 뜻하는 라틴어 Cohors 에서 유래
- 시간상 함께 진행되는 사람들의 집단 개념
- 과거: prospective, follow-up, longitudinal
현재: prospective 코호트 연구
retrospective / nonconcurrent 코호트 연구

전향적 및 후향적 코호트 연구



2. 세계의 역사적 변천

- Graunt and Halley : 생명표 개발
19세기 Farr : 인구집단 건강의 지표로 생명표 사용 촉진
- 1904년 보형산염(미국, 캐나다)
Specialized Mortality Investigation 발간
- 1900년 초 미국 결핵 요양소에서 전향적 코호트 연구 설계 사용
다양한 치료효과를 조사하기 위하여 생명표 방법 사용
1910년 영국 유사한 연구
- Frost : 코호트연구와 관련된 방법론을 진전시킴

(예1) Framingham Study

- 1940년 후반 일반 인구집단을 대상으로 심혈관질환의 위험요인 조사
↓
다른 질환들도 연구
↓
코호트 대상자의 자녀들 새로 등록 : 가족요인 제공
- 성공요인 : 적극, 협조가 잘되는 지역사회 선정
NIH의 지속된 지원
자료수집의 표준화된 프로토콜
- 다요인들의 장기자료에 적절한 다변량 통계방법이 없음
1967년 paper, discriminant analysis 적용

(예2) Investigation of the Atomic Bomb Survivors In Hiroshima and Nagasaki

- 전리방사선의 일회 노출 결과 조사
- 1945년 의학적 결과 조사
1946년 장기 연구 수행 결정
1947년 Atomic Bomb Casualty Commission 설립
방사선 노출량 재건, 생존자들의 코호트 연구
↓
급성 백혈병, solid tumors
방사선 표준의 토대
- 방법론적 이슈 축진 : 방사선 위험의 시간 및 연령 의존 방사선과 다른 요인들의 joint effect 측정 오류의 결과

(예3) US Public Health Service Study of
Colorado Plateau Uranium Miners

- 1950년대 초 고용기간의 지속적 노출
- 3,000명 이상이 1960년까지 등 록되어 현재까지 추적
누적노출과 노출률은 추적에 따라 다양
↓
라돈의 암발생위험 제시
시간에 따라 다양한 노출과 노출률에 대한 방법론적 진보
라돈과 흡연의 joint effects에 대한 방법개발 및 적용

(예4) Study of British Physicians

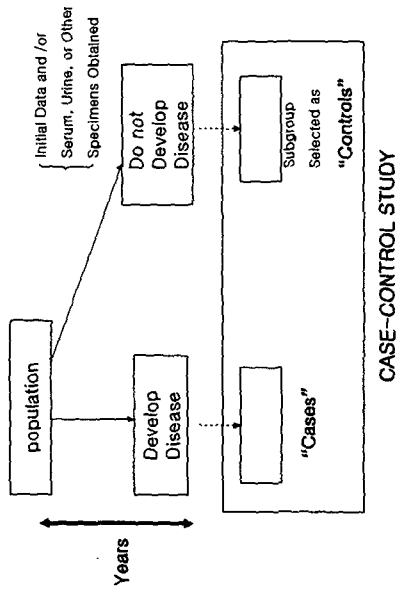
- Doll and Hill
- 환자-대조군 연구로 흡연과 폐암의 강한 연관성을 제시한 후
1951년 시작하여 현재까지
- 성공이유 : 사망추적이 수월한 협조적 연구집단 선정
주기적으로 흡연을 평가하는 간단한 우편 설문지 사용

- 1950년, 1960년대 수 많은 코호트 연구 정착
Washington County Study : 협정 저장 → 분자
역학
- 1970년대 규모가 큰 코호트 연구들
NIH 다기관 전향적 코호트연구(심혈관질환)
외적담당도 획득
표준화된 자료수집
Central coordinating center에서 자료축적,
평가 및 관리

- 자료결합 : 코호트 연구 수행이 용이
- 코호트 연구 자료분석의 복잡성 : 역학자와 생정보계학자의
협력
→로 방법론진보
- 설계에서는 노출과 결과의 관계를 효과적으로 평가하기 위
해
sampling methods 개
발

네스티드 환자-대조군 연구
네스티드 환자-코호트 연구

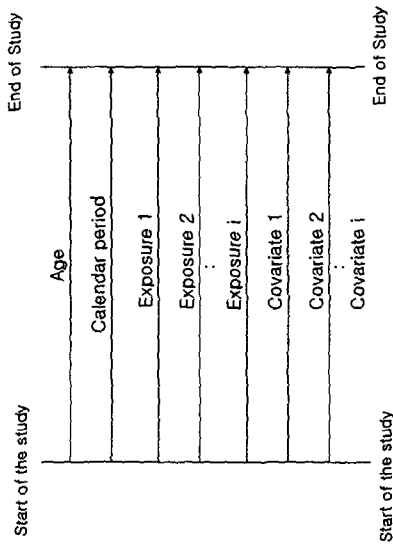
네스티드 환자-대조군 연구



제한점

- 추향적 설계 : 적절한 질의 역시적 자료가 유용 시 가능
- 전향적 설계 : 참여자의 적절한 추적이 유지될 때 가능
- 정보 바이어스
- 선택 바이어스
- 시간 특수바이어스에 대한 체계적 고려가 없음

다차원 시간



3. 분석방법의 역사적 변천

코호트 연구의 기본 결과 측정 분석방법의 진보

1. time-to-event
event-free time에 관한 분포 조사
marker에 대한 longitudinal data를 공변량으로 사용
2. repeated measures of markers of disease progression
event 발생이 marker의 변화에 미치는 영향 조사
time-to-event를 공변량으로 사용

이들간의 밀접한 관계를 하나의 동일한 framework로 설명하려는 시도가 많이
재
진행중

1970년대 이전

- life-table method
stratified approach (handling confounding, evaluating effect)
modification)
principal measure: binary variable (예: logistic regression)
- 1970년대
time-to-event data에 관한 regression method
(예: Cox's proportional hazard regression model)
events-in-person-years data에 관한 regression method

1980년대

- analysis of longitudinal data (경시적자료)에 관한 방법론 개발
- marginal model approach (주변모형 방법)
: multiple cross-sectional data 의 결합이 가능
 - transition model approach (전이모형 방법)
: past outcome 과 current outcome 간의 회귀모형
current/previous exposure와 current outcome 간의 회귀모형
 - random effect model approach (변량효과 모형 방법)
: 각 개인에 관한 unique 한 회귀 모수를 설정
개인들에 관한 average rate of change를 직접적으로 가
- 회귀
tree method for survival data

1990년대 전반

- generalized linear model (일반화 선형모형) 개발
(linear/logistic/Poisson/survival regression model)들을 unified framework으로 설명
계산방법 : quasi-likelihood method(의사 가능도 방법)의 개발
late entries into observation에 관한 분석방법 개발
underlying hazard에 관한 direct measure 방법 개발
(parametric survival method에 다시 관심을 기울이기 시작함)
(참고: Cox의 PH model은 relative hazard를 계산하는 방법임)
- nested design에 관한 분석방법 개발
(예: conditional logistic regression method for nested case-control study; Cox's regression method for nested case-cohort study 등)

Overview of analytical methods for cohort studies

Outcome	Comparison		Measure of Association
	Summary measure	Exposed/unexposed (2-sample)	
Events in person-years Time to event	Incidence rate Kaplan-Meier/maximum likelihood estimates	(O-E)/var Logrank or Mantel-Haenszel/health ratio test Extended logrank	Poisson Proportional hazards/parametric Relative incidence Relative hazard/relative percentile or line Relative hazard
Time to event: exposures matched case-cohort	Extended Kaplan-Meier Proportion exposed	Propitied chi-square or McNemar (Robust) logrank	Proportional hazards, staggered entries Conditional logistic Odds ratio
Intermediate outcome repeatedly measured	Proportion exposed Change	Regression for correlated data: marginal, conditional, random effects	Relative hazard Differences in change over time

Software

- SAS (SAS Institute, Cary, North Carolina)
- Splus (Statistical Sciences, Inc., Seattle, Washington)
- STATA (Stata Corporation, College Station, Texas)
- EGRET (Cytel Software Corporation, Cambridge, Massachusetts)