

연구설계 심사를 위한 리뷰팁

김현창

연세대학교 의과대학 예방의학교실 교수

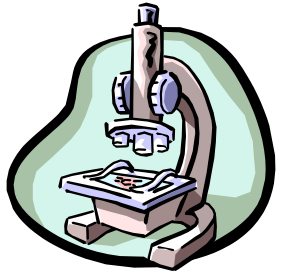
연세대학교 디지털헬스케어혁신연구소장

Editor-in-chief, Epidemiology & Health

연구설계 (Study Design)

- 연구목적을 달성하기 위한 구체적 계획으로 자료원, 자료수집, 자료처리, 자료분석, 결과해석 등을 자 수집 전에 계획하는 연구의 틀을 말한다.

Three Approaches to Health



Basic Science



Clinical Science



Population Science

역학연구 설계의 분류: 세가지 기준으로 나누어 보자

(1) 가설 유무(연구목적)에 따라

- 기술
- 분석

(2) 개입(중재) 여부에 따라

- 실험 연구
- 비실험(관찰) 연구

(3) 시간적 순서에 따라

- 단면적 연구
- 전향적 연구
- 후향적 연구

역학적 연구 방법의 분류

(1) 가설(Hypothesis) 유무에 따라

• 기술 (Descriptive) 연구

- 사례 연구 (Case study), 사례군 연구 (Case series study)
- 생태학적 연구 (Ecologic study)
- 단면 연구 (Cross-sectional study)

• 분석 (Analytic) 연구

- 실험 연구 (Experimental study)
- 코호트 연구 (Cohort study)
- 환자대조군 연구 (Case-control study)
- 단면 연구 (Cross-sectional study)
- 생태학적 연구 (Ecologic study)

가설(Hypothesis, 假說)

- 아직 검증되지 않았으나 검증할 가치가 있는 설(說), 주장
- 연구할 수 있는 과학적 의문 (Question, 疑問)

기술역학 vs. 분석역학

역학: 인구집단(다수의 사람들)에서 질병(건강)의 분포와 결정요인을 밝히는 학문

1. Descriptive epidemiology (기술역학)

- Study of the occurrence and distribution of disease
- Distribution of disease by **TIME, PLACE, and PERSON**

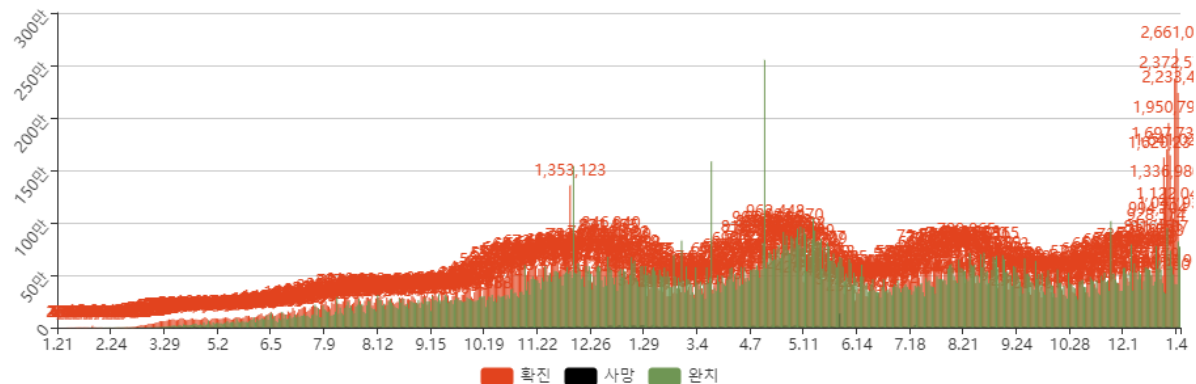
2. Analytic epidemiology (분석역학)

- Study of the **CAUSES** and **RISK FACTORS** of disease
- Relationships between a risk factor and a disease
- Strategies for disease prevention and treatment
- At individual and community levels

Descriptive epidemiology (기술역학) - 감염병

전세계 코로나19 (COVID-19) 추이

총 확진자: 298,200,216명, 사망: 5,483,731명, 완치: 255,102,404명



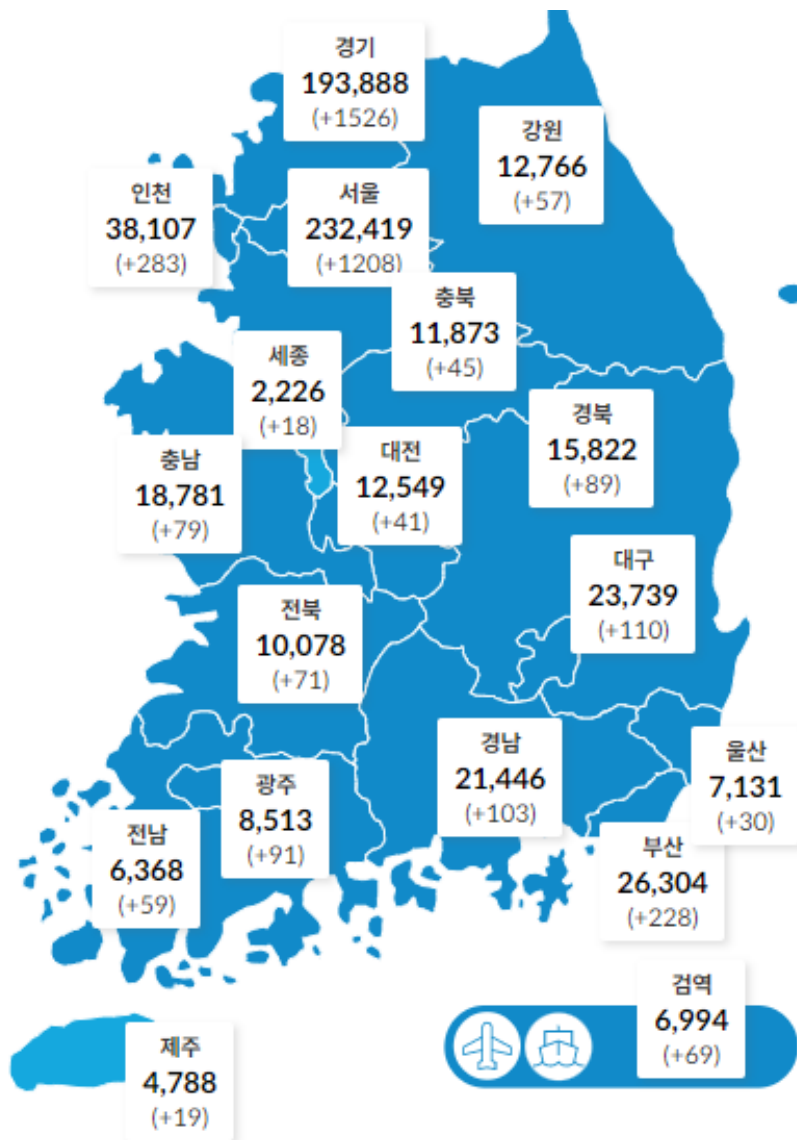
확진자 성별 현황 (1.600시 기준)

구분	확진자(%)	사망자(%)	치명률(%)
남성	339,743 (51.96)	3,035 (51.55)	0.89
여성	314,049 (48.04)	2,852 (48.45)	0.91

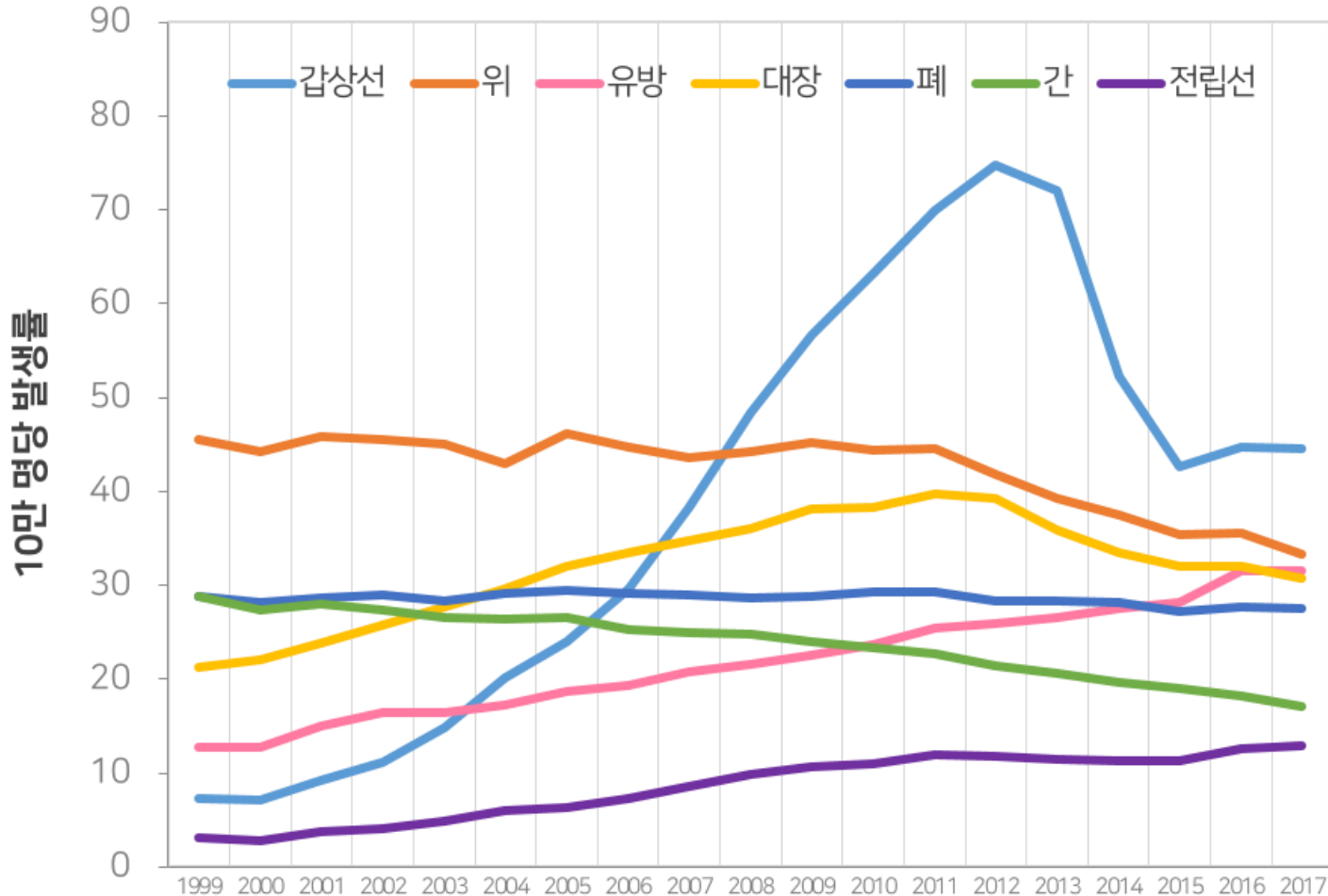
* 치명률 = 사망자수 / 확진자수 × 100

확진자 연령별 현황 (1.600시 기준)

구분	확진자(%)	사망자(%)	치명률(%)
80 이상	21,018 (3.21)	2,935 (49.86)	13.96
70-79	39,283 (6.01)	1,610 (27.35)	4.10
60-69	92,411 (14.13)	923 (15.68)	1.00
50-59	94,469 (14.45)	284 (4.82)	0.30
40-49	95,799 (14.65)	81 (1.38)	0.08
30-39	95,262 (14.57)	38 (0.65)	0.04
20-29	97,431 (14.90)	13 (0.22)	0.01
10-19	65,794 (10.06)	0 (0.00)	0.00
0-9	52,325 (8.00)	3 (0.05)	0.01



주요 암종별 연령표준화발생률 추이
남녀 전체, 1999-2017



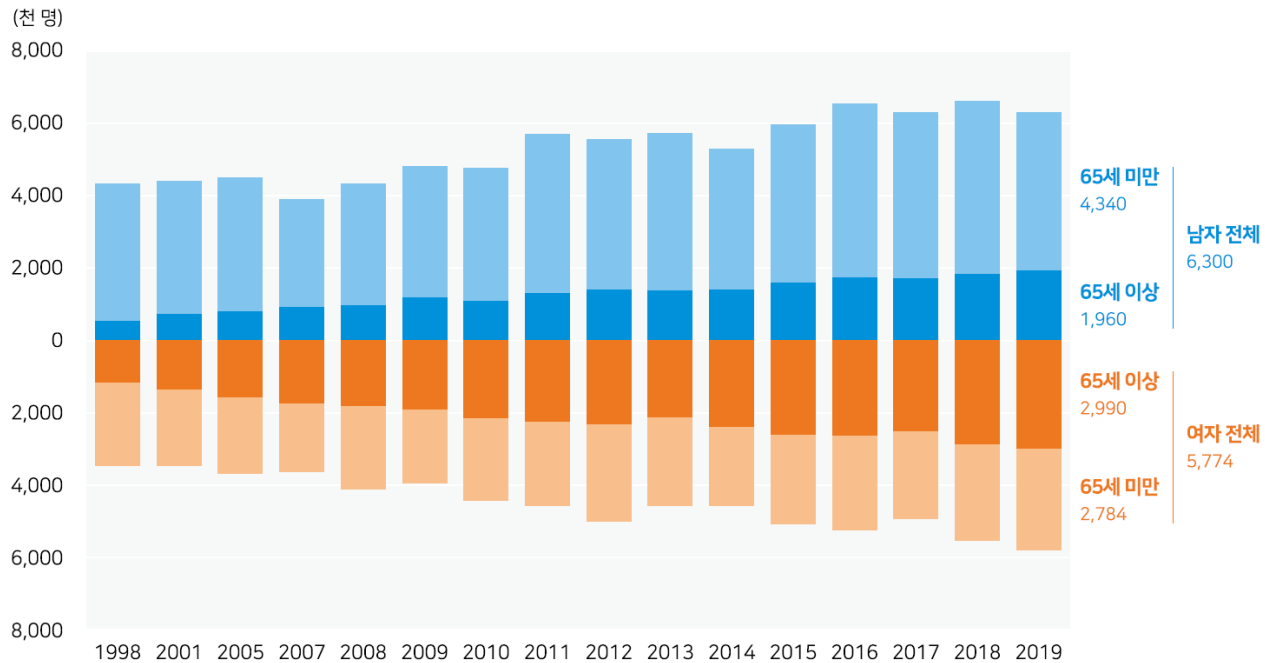
연령표준화발생률: 우리나라 2000년 주민등록연앙인구를 표준인구로 사용

암종	추이 1		추이 2	
	발생 기간	연간% 변화율	발생 기간	연간% 변화율
갑상선	1999-2011	22.5*	2011-2017	-10.8*
위	1999-2011	-0.2	2011-2017	-4.7*
유방	1999-2002	9.8*	2002-2017	4.6*
대장	1999-2010	5.9*	2010-2017	-4.1*
폐	1999-2010	0.2	2010-2017	-0.9*
간	1999-2010	-1.7*	2010-2017	-4.5*
전립선	1999-2009	14.3*	2009-2017	1.6*

* p < 0.05

고혈압 추정 유병자수 변화

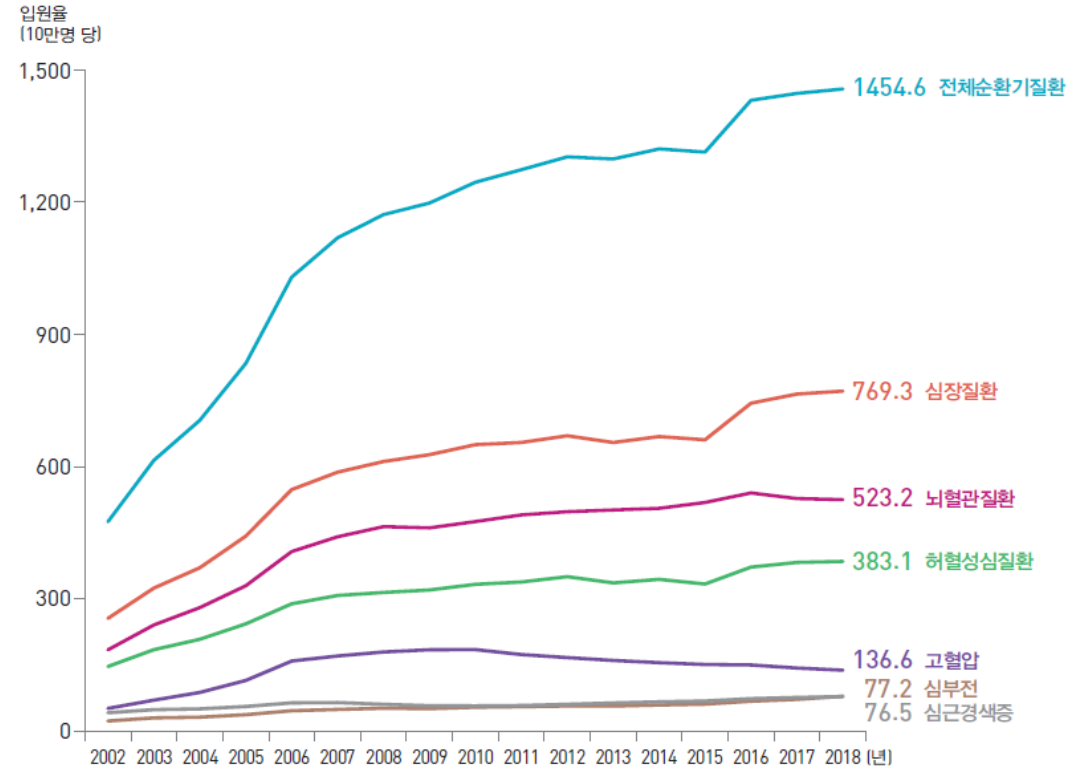
(20세 이상)



자료원: 국민건강영양조사 1998-2019

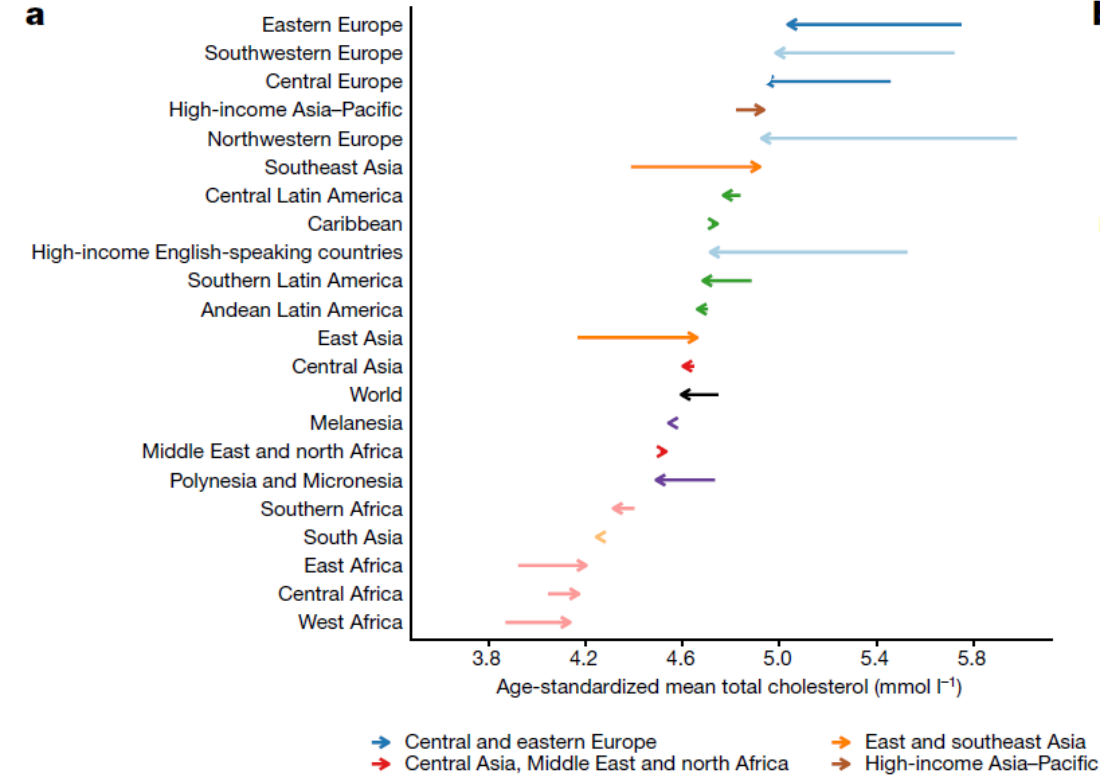
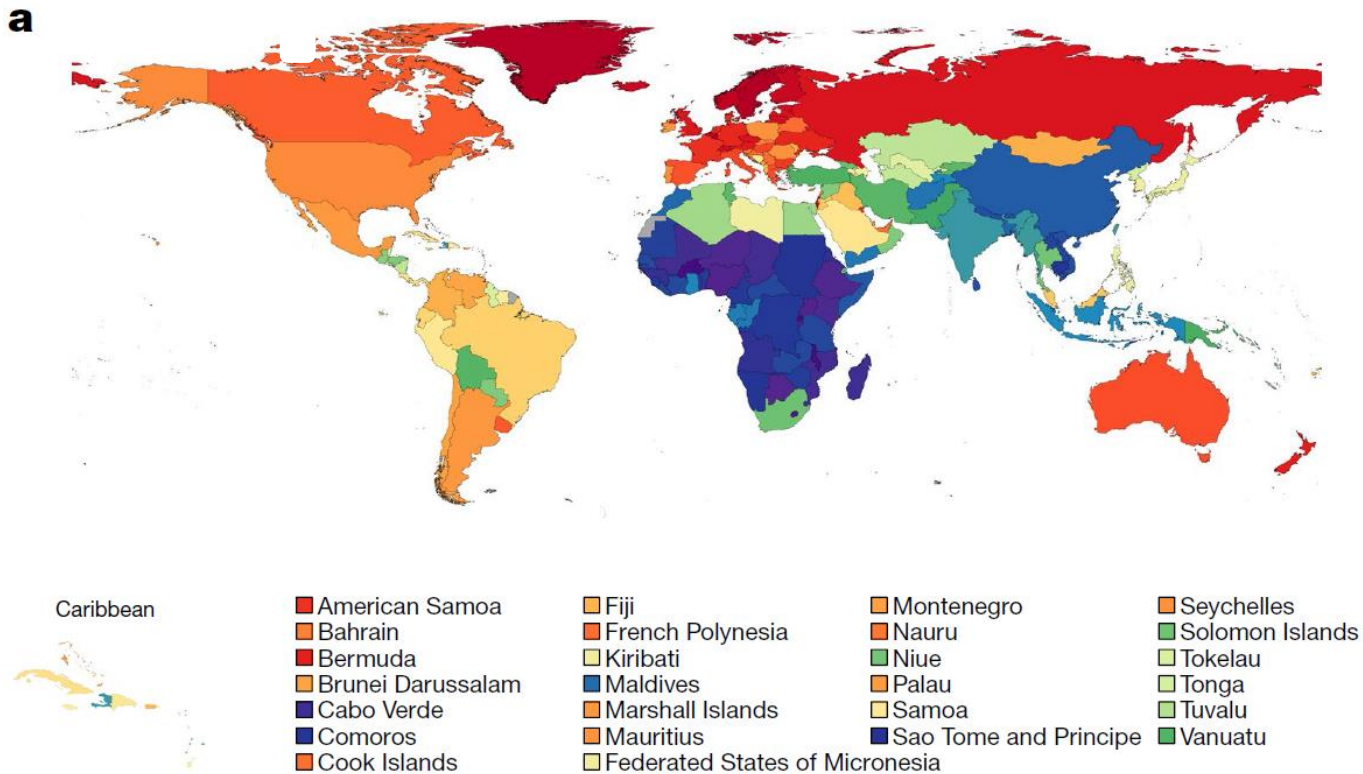
전체순환기질환 및 주요 심혈관질환 입원율 변화

〈주상병 기준〉



- 주상병명을 기준으로 하면, 2002-2018년 사이 10만명당 입원율은 전체순환기질환은 3.1배 증가 (474 → 1455), 심장질환은 3.0배 증가 (255 → 769), 허혈성심질환은 2.6배 증가 (146 → 383), 심근경색증은 1.9배 증가 (41 → 77), 심부전은 3.5배 증가 (22 → 77), 고혈압은 2.7배 증가 (50 → 137), 뇌혈관질환은 2.9배 증가 (183 → 523) 하였음

Descriptive epidemiology (기술역학) - 국제공동연구



NCD Risk Factor Collaboration (NCD-RisC). "Repositioning of the global epicentre of non-optimal cholesterol." *Nature* 582 (2020): 73-77.

기술역학 vs. 분석역학

역학: 인구집단(다수의 사람들)에서 질병(건강)의 분포와 결정요인을 밝히는 학문

1. Descriptive epidemiology (기술역학)

- Study of the occurrence and distribution of disease
- Distribution of disease by **TIME, PLACE, and PERSON**

2. Analytic epidemiology (분석역학)

- Study of the **CAUSES** and **RISK FACTORS** of disease
- Relationships between a risk factor and a disease
- Strategies for disease prevention and treatment
- At individual and community levels

역학적 연구 방법의 분류

(2) 개입(Intervention) 여부에 따라

- 실험연구(Experimental study), 중재(Intervention study)
- 관찰 연구(Observational Study)
 - 코호트 연구 (Cohort study)
 - 환자대조군 연구 (Case-control study)
 - 단면 연구 (Cross-sectional study)
 - 기술연구 (Descriptive epidemiology)

역학적 연구방법의 분류



역학적 연구방법의 분류

(3) 원인-결과 확인의 시간적 순서(temporal relationship)에 따라

- **전향적 연구**

- 코호트 연구 (Cohort study)
- 실험 연구 (Experimental study)

- **후향적 연구**

- 환자대조군 연구 (Case-control study)

- **단면적 연구**

- 단면 연구 (Cross-sectional study)

분석역학에 쓰이는 대표적 연구방법들

무작위대조시험 (randomized controlled trial, RCT)

- 원인(위험요인) 노출 여부를 연구자가 무작위 배정하고, 그에 따른 결과(질병) 발생빈도를 비교함

코호트 연구 (cohort study)

- 다수의 사람들에서 원인(위험요인)을 측정해 두고, 시간이 지나면서 결과(질병) 발생을 확인함

환자-대조군 연구 (case-control study)

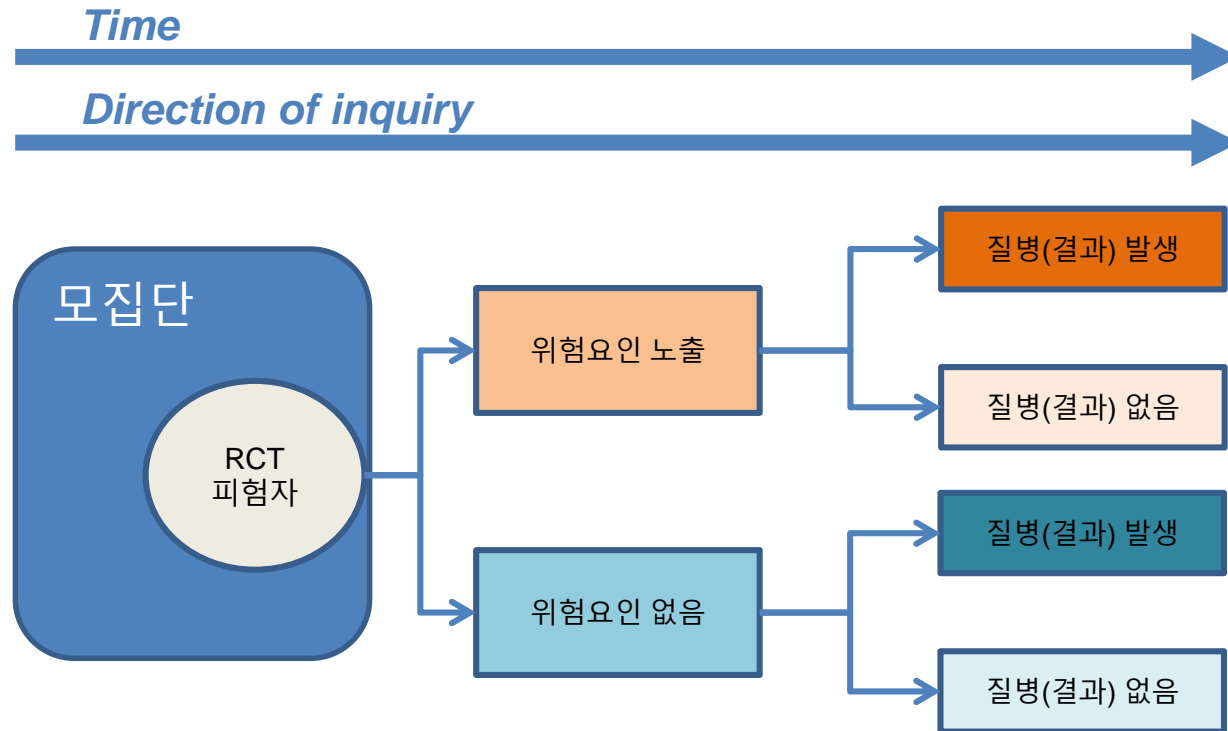
- 결과(질병) 유무에 따라 사람들을 먼저 분류하고, 원인을 나중에 확인함

단면연구 (cross-sectional study)

- 다수의 사람들에서 원인(위험요인)과 결과(질병)를 동시에 측정함

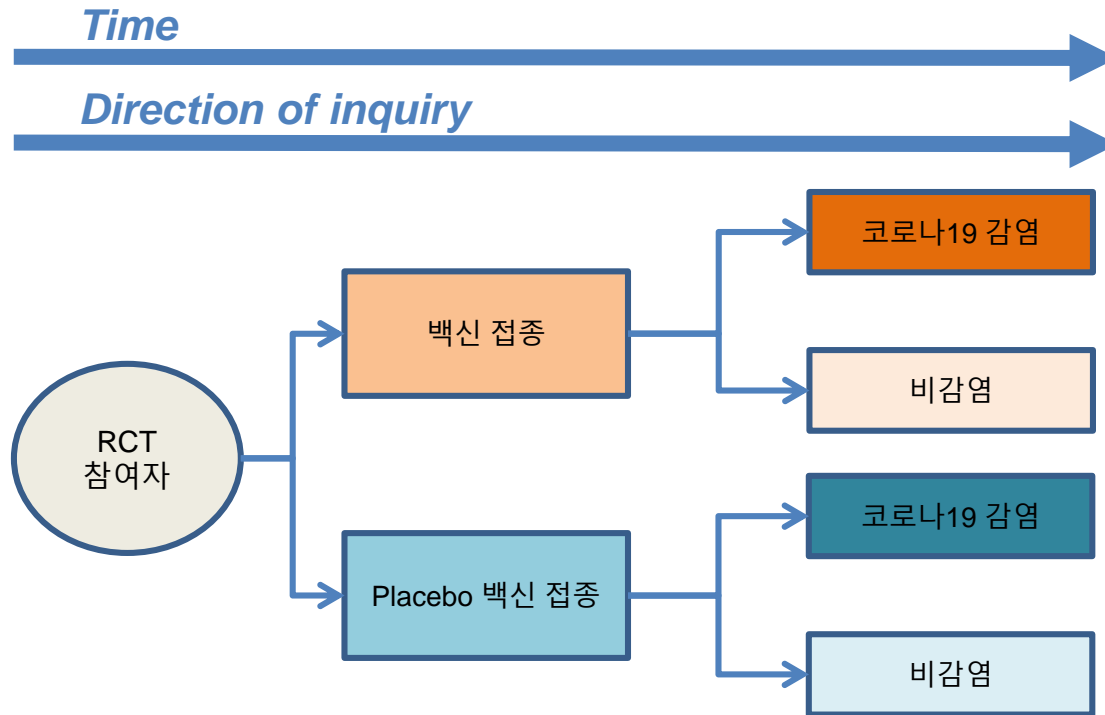
Randomized Controlled Trial (RCT)

- 모집단에서 피험자를 모집하고 위험요인(원인) 노출 여부를 연구자가 무작위로 배정한 후에
- 시간이 지나면서 질병(결과) 발생을 관찰하여, 위험요인에 따른 질병 발생 빈도를 비교하는 연구



Randomized Controlled Trial (RCT)

코로나19 백신을 맞으면 감염율이 낮아질까?



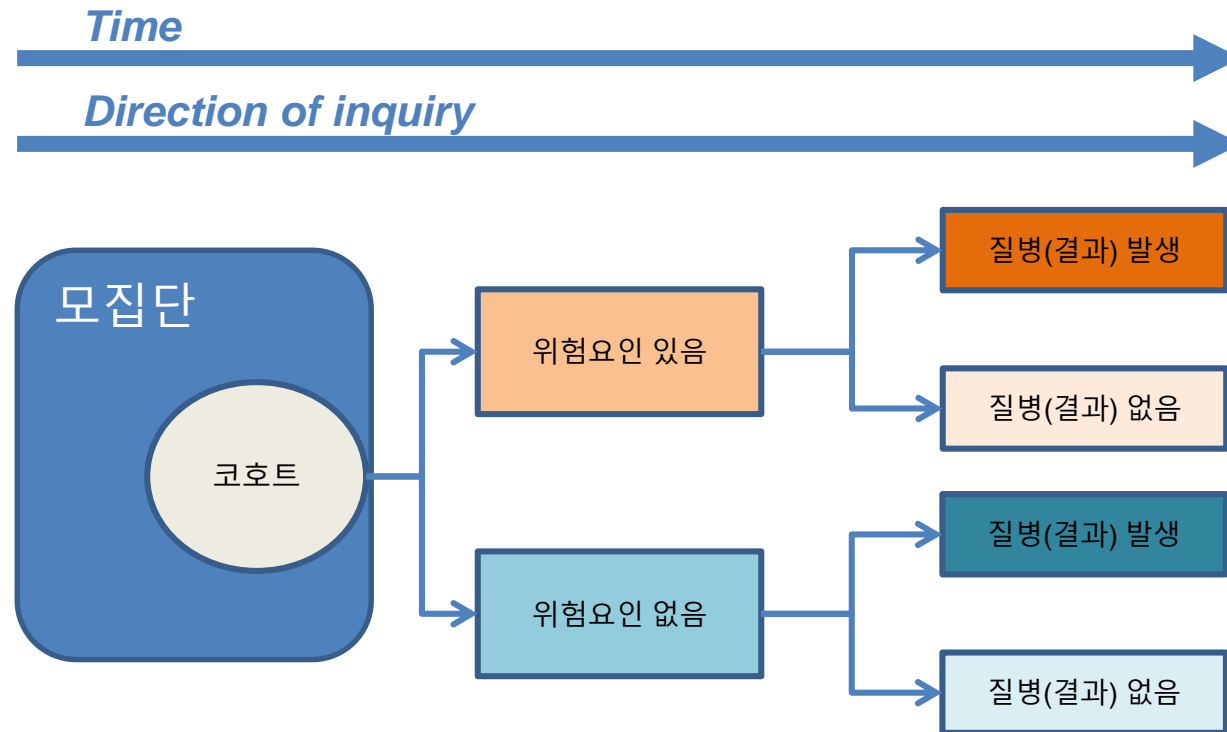
	비감염	감염
백신접종 (100,000명)	99,990	10
Placebo 백신 (100,000명)	99,900	100

$$\text{비교위험도} = \frac{\left(\frac{10}{100,000}\right)}{\left(\frac{100}{100,000}\right)} = 0.10$$

$$\text{백신효능} = \frac{\left(\frac{100}{100,000} - \frac{10}{100,000}\right)}{\left(\frac{100}{100,000}\right)} = 0.90 = 90\%$$

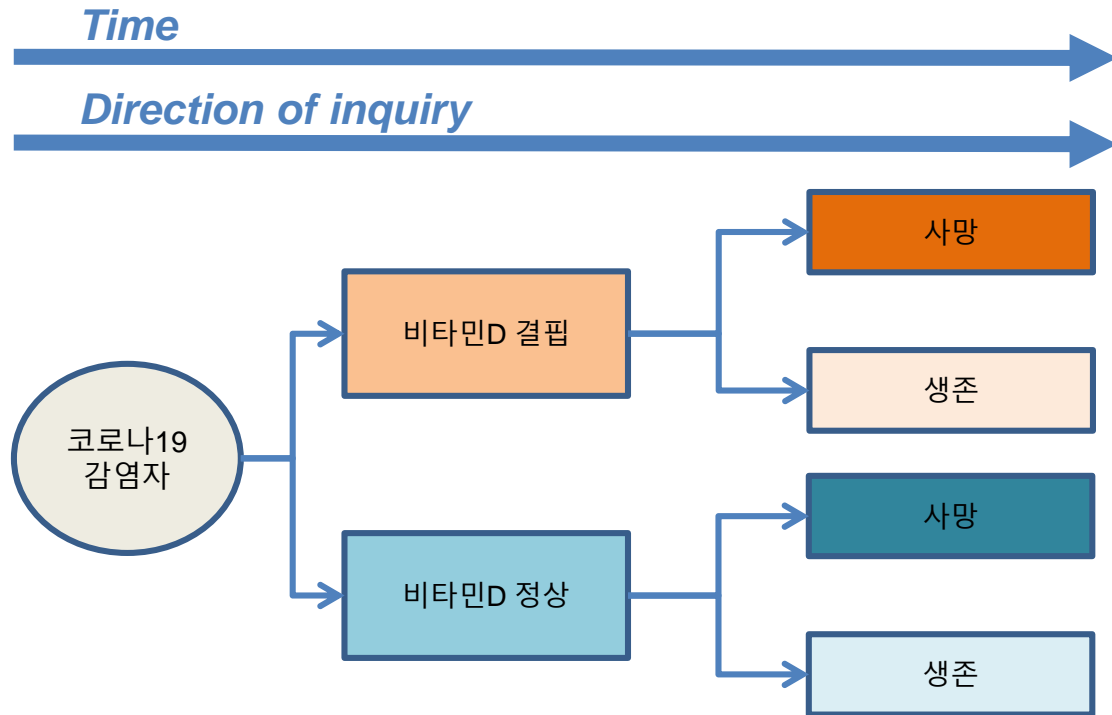
Cohort study (코호트 연구)

- 모집단에서 코호트를 추출하고 위험요인(원인) 노출 여부를 먼저 측정해 놓은 후에
- 시간이 지나면서 질병(결과) 발생을 관찰하여, 위험요인에 따른 질병 발생 빈도를 비교하는 연구



Cohort study (코호트 연구)

코로나19 감염자 중에서 비타민D 결핍이 있으면 중증으로 사망할 위험이 더 높을까?

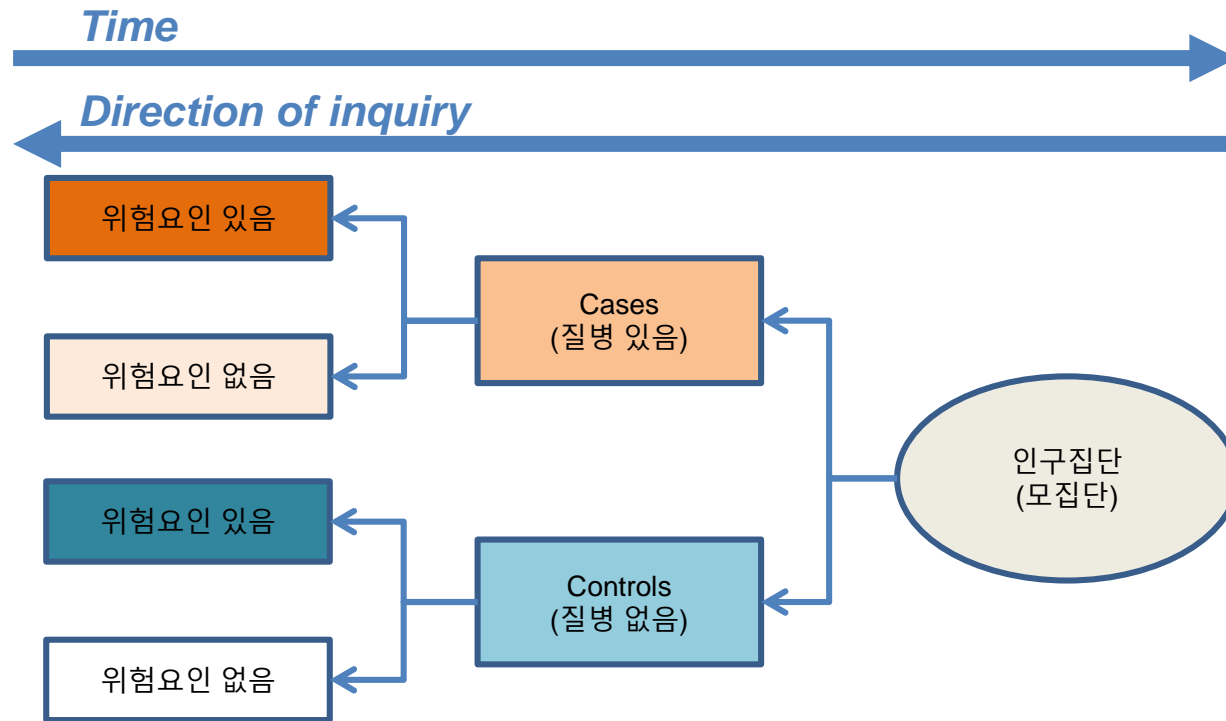


	생존	사망
비타민D 결핍 (1,000명)	998	2
비타민D 정상 (1,000명)	999	1

$$\text{비교위험도} = \frac{\left(\frac{2}{1,000}\right)}{\left(\frac{1}{1,000}\right)} = 2.00$$

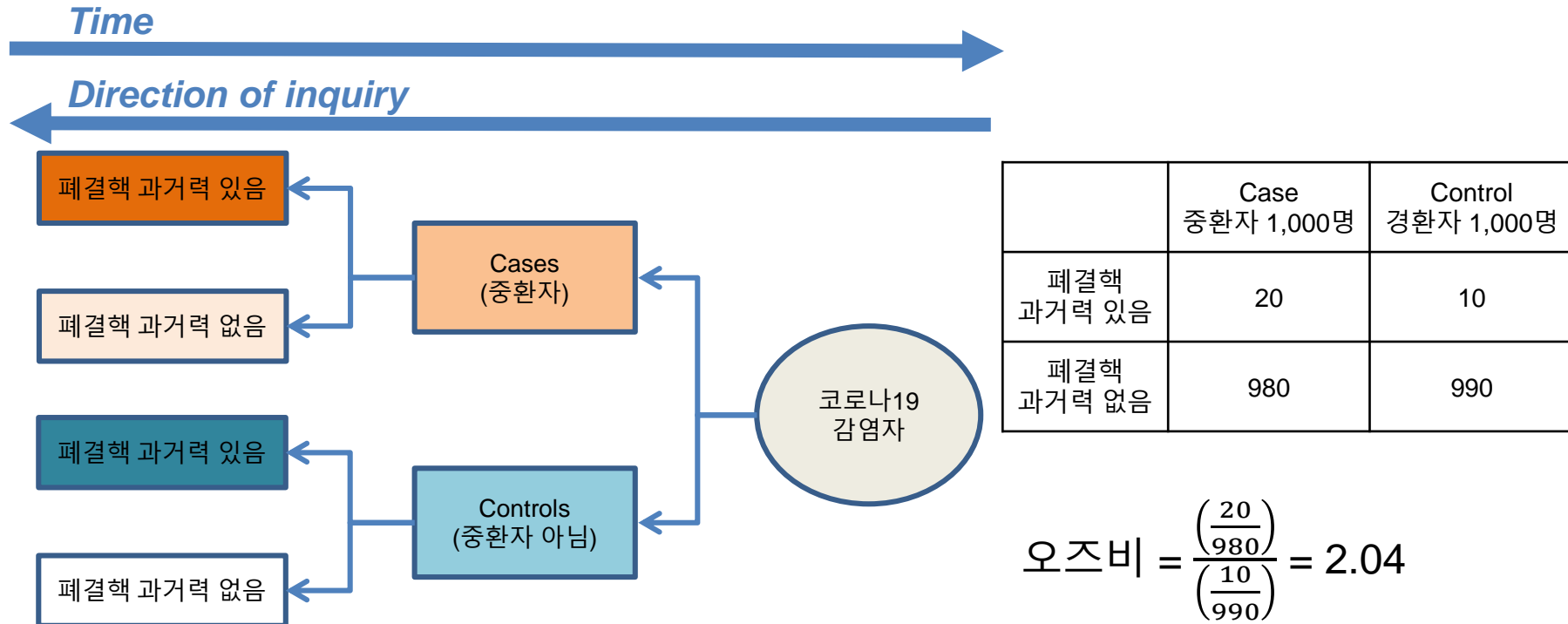
Case-control study (환자-대조군 연구)

- 모집단에서 환자군(질병 있는 사람)과 대조군(질병 없는 사람)을 먼저 추출한 후에,
- 과거에 위험요인(원인)에 노출 되었는지를 파악하여 위험요인과 질병의 관련성을 평가하는 연구



Case-control study (환자-대조군 연구)

코로나19 감염자 중에서 과거 폐결핵을 앓았던 사람은 중증으로 이환될 가능성이 높을까?



Cross-sectional study (단면연구)

- 특정시점(또는 짧은 기간)에 인구집단을 대상으로 위험요인(원인)과 질병(결과)여부를 조사하여
- 위험요인과 질병의 관련성을 평가하는 연구

	코로나19 (+)	코로나19 (-)
흡연	5	330
비흡연	5	660

코로나 양성률 0.1%

흡연율 33.5%

흡연자중 코로나 양성 1.49%

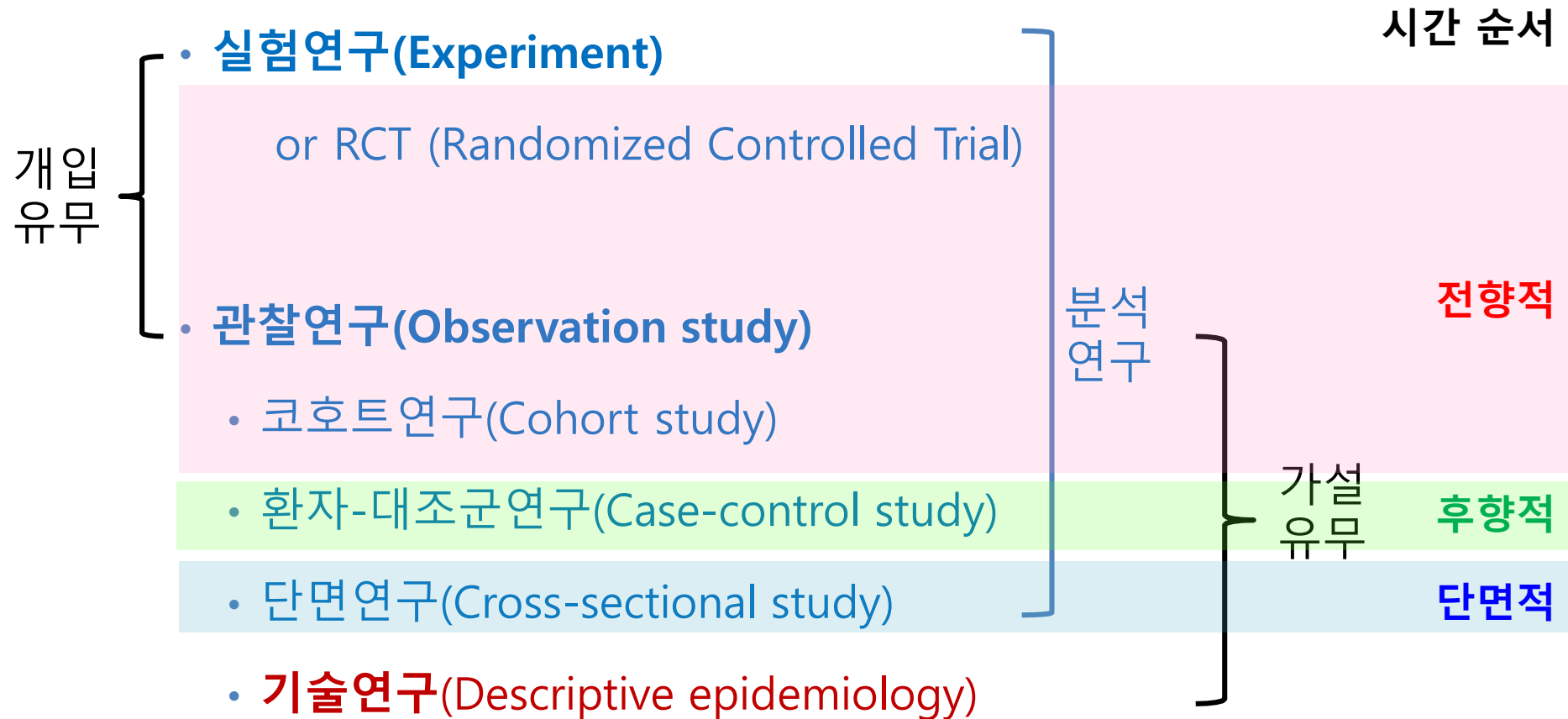
비흡연자중 코로나 양성 0.75%

코로나 양성자 중 흡연 50%

코로나 음성자 중 흡연 33%



역학적 연구방법의 분류



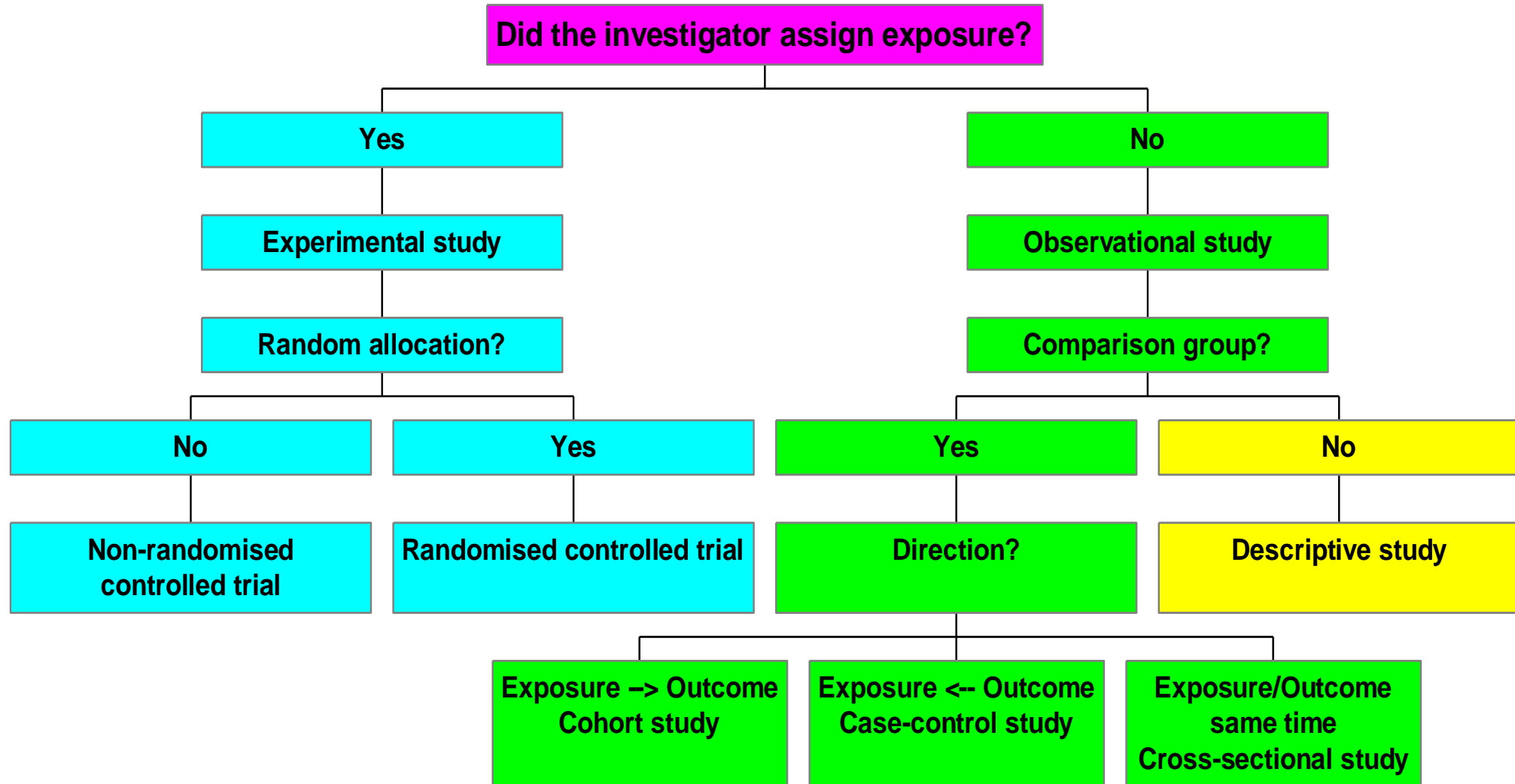
주요연구설계의장·단점

무작위대조시험(RCT)	코호트 연구	환자-대조군 연구	단면 연구
<ul style="list-style-type: none"> 가장 결정적 근거 제공 Confounding의 영향을 효과적 통제 원인-결과 선후 관계 명확함 	<ul style="list-style-type: none"> 원인-결과 선후 관계 명확함 Confounding의 영향 통제 가능함 원인-결과 선후 관계 명확함 동시에 여러 질병(결과) 연구가 가능함 흔치 않은 원인(위험요인) 연구에 유리함 	<ul style="list-style-type: none"> 흔치 않은 질병(결과) 연구에 유리함 동시에 여러 원인(위험요인) 연구가 가능함 시간, 비용이 적게 필요 윤리적 문제 적음 	<ul style="list-style-type: none"> 시간, 비용이 적게 필요 윤리적 문제 적음 위험요인에 따른 질병 확률 계산, 질병 여부에 따른 위험요인노출 확률 계산 양방향 연구가 모두 가능 기술 역학 연구에 적용 가능
<ul style="list-style-type: none"> 윤리적 문제 제한된 주제만 가능 시간, 비용 많이 필요 	<ul style="list-style-type: none"> 시간, 비용 많이 필요 윤리적 문제 약간 있음 	<ul style="list-style-type: none"> 원인-결과 선후 관계 불분명 Recall Bias에 취약함 적절한 대조군 찾기 어려움 	<ul style="list-style-type: none"> 원인-결과 선후 관계 불분명 Confounding의 영향을 많이 받음 원인-질병 둘다 드물지 않은 경우에 적합

연구설계의 세가지 분류기준에 따른 장단점이 정해진다

- 연구목적에 따라: 기술 연구(가설 없음) vs. 분석 연구(가설 있음)
- 개입여부에 따라: 중재(실험, 개입) 연구 vs. 관찰 연구
- 시간적 순서에 따라: 단면적 연구 vs. 전향적 연구 vs. 후향적 연구

이미 수행된 연구의 연구설계 파악



새로운 or 하이브리드 디자인

Pragmatic Trial

Mendelian Randomization Study

Case-crossover study

Case-cohort study

Nested case-control study

IPD (individual participant data) meta-analysis

연구설계라고 보기는 어렵지만...

Propensity Score (PS) matching

Inverse Probability Treatment Weighting (IPTW)

Common Data Model (CDM)

Big Data Analysis

Real-world data (RWD) or Real world evidence (RWE)

Electronic Medical Record (EMR) analysis

Insurance Claims Data analysis

연구 설계의 구분: 상대적인 개념으로 이해하자

구분	차이점
Qualitative vs. Quantitative	수치 혹은 객관적 측정 가능한 데이터
Descriptive vs. Analytical	가설 검정, 비교대상, 통계검정
Exploratory vs. Confirmatory	사전 정의된 가설 검정
Observational vs. Experimental	중재 여부
Prospective vs. Retrospective	(1)원인과 결과 측정의 선후관계 (2)연구가설과 데이터수집의 선후관계
Explanatory vs. Pragmatic trial	실제 의료 환경에서 효과를 검증
Primary vs. Secondary data analysis	연구목적을 위해서 수집된 데이터
Primary research vs. Meta-research	연구를 대상으로 하는 연구