

연구방법

; 독자를 어떻게 설득할 것인가?

KAMJE

Korean Association of
Medical Journal Editors

고상백

연세대학교 원주의과대학

들어가며; 논문이란?

□ 논문이란 특정한 구체적인 주제에 대하여 연구자가 자신의 의견 및 주장이 포함된 학문적 연구결과를 일정한 형식과 논리적 과정에 맞추어 체계적으로 작성한 글

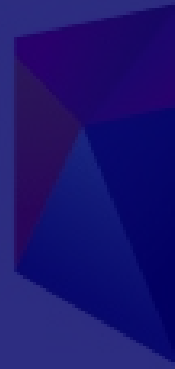
□ 논문의 요건

; 구체성, 창의성, 검증 및 재현 가능성, 효용성 및 유용성 등

1. 연구방법을 기술하는 이유

- 연구방법이 목적달성에 적합해야 함
- 재현가능성, 검증가능성
- 정확하고 자세하게 기술
- 객관적 서술
- 체계적 서술

KAMJE
Korean Association of
Medical Journal Editors



2. 독자는 무엇을 보고 싶은가?

□ 이 연구결과를 믿을 수 있는가?

– 누구를 연구했는가?

; 일부 병원의 환자인가? 지역의 표본인가?

건강보험공단 자료인가? 코호트 자료인가?

– 이 방법으로 할때 결과를 신뢰할 수 있나? 일반화 가능한가?

□ 이 연구질문에 왜 이 연구설계를 사용했는가?

– 연구질문과 연구설계의 적합성

KAMJE
Korean Association of
Medical Journal Editors

2. 독자는 무엇을 보고 싶은가?

□ 내가 이 논문의 방법대로 하면 나도 할 수 있을까?

- 좋은 Methods는 누가 읽어도 동일하게 재현할 수 있어야 한다.
; 변수정의가 명확한가? 구체적으로 제시되어 있는가?
재현가능한가?

□ 바이어스가 최소화 되었는가? 이 결과가 틀릴 가능성은 없는가?

- 선택바이어스, 정보바이어스, 혼란 등의 통제방법 제시

□ 통계분석이 적절한가?

- 연구질문에 맞는 통계를 사용했는가?

3. 연구방법의 기술순서

□ 학술지에 따라 다르지만 일반적으로 아래 사항들이 순서에

따라 기술하도록 권고

- 연구모형의 설계, 연구대상자, 변수의 선정,
연구자료 또는 재료, 연구기간, 분석방법 등

KAMJE
Korean Association of
Medical Journal Editors

4. 연구 설계에 따른 대상자 선정 방법

□ 연구대상자 정확하게 선정하려면 연구설계에 대한 명확한 제시가 중요

기술연구

집단단위 연구(상관연구)

개인단위 연구

사례보고

사례군 보고

단면연구

분석연구
Kor Med Assoc of
Medi Journal Editors

관찰

환자-대조군 연구

코호트연구

실험

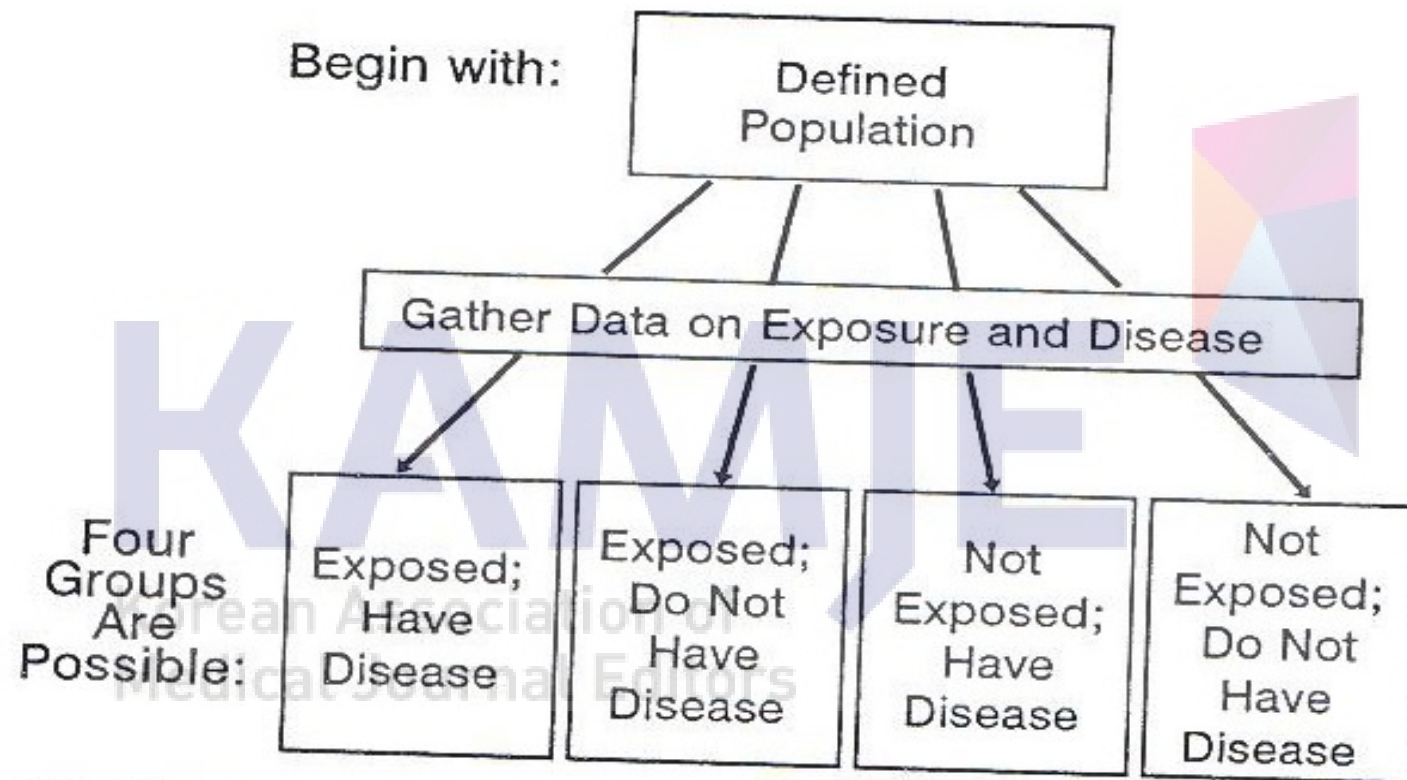


FIGURE 10-10 ▼ Design of a cross-sectional study: I.

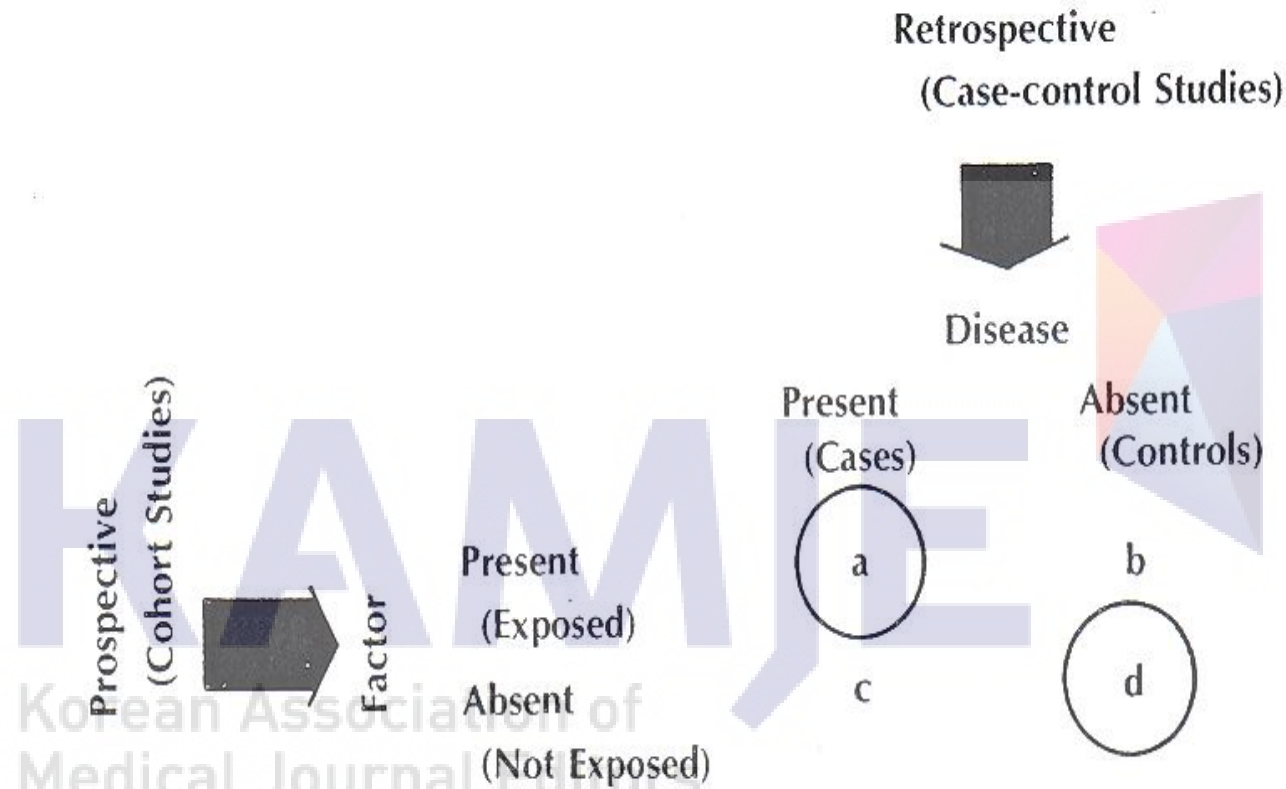


Figure 7-3 Fourfold table (2 × 2) showing retrospective and prospective study design.

4-1. 환자대조군 연구의 특징

□ 가설 검증을 위해 의학연구에서 초기 단계에 가장 널리

□ 기본개념

– 연구하고자 하는 질병이 있는 집단(환자군)과 없는 집단(대조군)을 선정하여 질병의 발생과 관련되어 있으리라 생각하는 잠재적 위험요인에 대한 두 집단의 과거 노출율을 비교하는 방법

– **비교성**이 핵심

; 환자군의 정확한 정의와 대조군 선정 방법

환자군의 정의 (Definition of Cases)

- 자격기준(Eligibility Criteria)을 수립하는 이유:
연구대상을 위험요인에 노출될 가능성이 있는 집단에 한정하여 선정하기 위한 것.

예) 최근의 피임약 복용과 심근경색증과의 관계 연구 시 제외 대상자:
폐경기 이후 부인
불임수술을 한 부인
피임약을 복용할 수 없는 당뇨병 등
만성질환을 가진 부인
- 환자군의 정의 ; 정확한 진단명
환자군을 모두 동일한 질병으로 구성하려면 진단기준이 명확해야 함

대조군의 정의 (Definition of Controls)

- 자격기준(Eligibility Criteria)

위험요인에 노출될 가능성이 있는 집단에 한정하여야 함.

특히 병원중심 환자-대조군연구는 연구하려는 위험요인과 관련된 질환을 가진 사람도 제외하여야 함.

예) 아스피린 복용과 심근경색증과의 관계에 대한 연구에서, 만성 관절염이나 소화성 궤양 환자는 아스피린을 복용하였을 가능성이 높으므로 대조군으로서의 자격이 없다.

- 대조군의 정의(Definition of Controls)

비교성(comparability)이 높아야 한다.

그러나 일반적으로 환자군과 대조군이 완벽한 비교성을 갖추기는 어렵다. 이러한 문제점은 분석 시에 부분적으로 보완할 수 있다

- 대조군 선정 및 추출방법

; 짝 추출, 도수대응추출 등

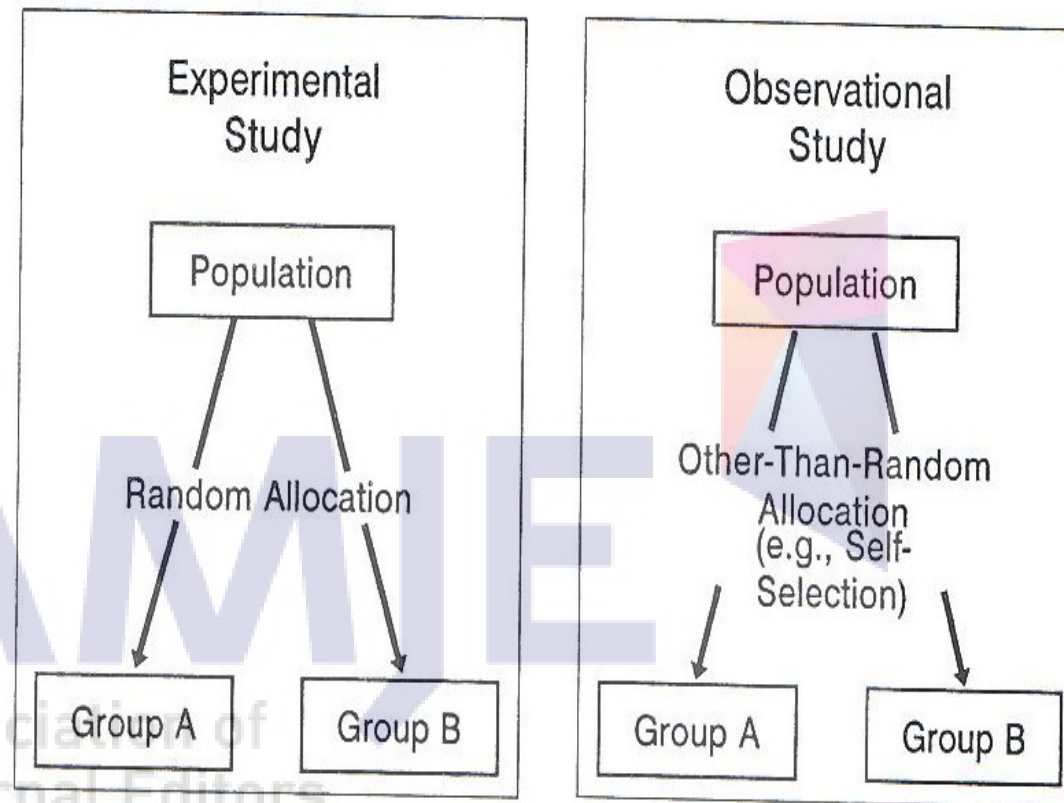
4-2 코호트연구(Cohort Studies)

- 질병의 원인과 관련되어 있다고 생각되는 어떤 요소를 가진 집단과
 갖지 않은 집단을 계속 관찰하여 두 집단의 질병발생률 혹은
사망률을
 비교하는 연구방법.
- 노출집단의 선정하기에 편리한 집단
 - 첫째, 연구하려는 요인에 흔히 노출되거나 경험을 갖고 있는 집단.
 예) 특수 직업군이나 작업환경에 노출된 집단.
 - 둘째, 노출정도나 질병 발생을 쉽게 확인할 수 있는 체계를 갖춘
 집단.
 예) 의료보험이나 생명보험 가입자 및 의사(醫師) 집단
 - 셋째, 장기간의 추적조사가 가능한 집단.
 - 넷째, 질병 발생률이 높은 집단

4-3 실험 연구

- 처치군과 비교군을 추적 관찰하여 처치(intervention)의 효과를 비교하는 역학적 연구
- 연구대상자를 장기간 추적 관찰한다는 점에서 코호트연구와 비슷하나 관찰과는 달리 연구자의 처치가 개입된다는 점이 다르다.
- 예) Coronary Primary Prevention Trial(CPPT)에서는 혈중 콜레스테롤치가 높은 남자 3,806명을 두 집단으로 나누어 한쪽 집단에는 혈중 콜레스테롤치를 낮추는 약인 콜레스티라민을 투여하고, 다른 집단에는 위약(placebo)을 투여하여 7-10년간 추구 관찰한 결과, 콜레스티라민 투여군의 심혈관질환 사망률이 유의하게 낮았다.
- 실험은 인과관계를 설정하는 데 가장 결정적인 정보를 제공한다.

FIGURE 9-3 ▼ Selection of study groups in experimental and observational epidemiologic studies.



5. 변수의 선정과 측정 방법

- 가설을 검증하기 위한 독립변수와 종속변수에 대한 정의와 측정방법을 명확하게 구체적으로 제시
- 설문조사의 경우 신뢰도와 타당도가 검증된 출처를 제시하거나, 새로운 방법의 경우 연구자가 신뢰도와 타당도 결과를 제시
- 진단도구의 경우도 신뢰도와 타당도를 제시

5-1. 측정방법의 신뢰도와 타당도

구분	신뢰도	타당도
정의	동일 대상의 반복 측정 시 그 값이 비슷한 정도	측정하고자 하는 것을 바르게 측정한 정도
측정방법	반복 측정한 값들을 비교	측정값을 비교값(참값)과 비교
중요성	통계학적 검정력 증가	결론의 타당성 증가
저하 요인	확률오차	체계오차

신뢰도를 높이는 방법

첫째, 측정방법을 표준화한다.

둘째, 조사원을 훈련시킨다.

셋째, 정밀한 측정도구를 사용한다.

기계식보다는 전자식 도구의 신뢰성이 높다.

넷째, 자동화된 도구를 사용한다.

다섯째, 같은 변수를 여러 번 측정하여 그 평균값을 사용한다.

여섯째, 표본수를 증가시킨다.

일곱째, 신뢰성이 높은 표본추출방법을 사용한다.

여덟째, 효율성이 높은 통계학적 분석법을 사용한다.

내적 타당도를 향상시키기 위한 방법

앞에서 설명한 신뢰도를 향상시키는 방법 중 첫 네 가지 외에

다섯째, 신중한 측정을 하여야 한다.

여섯째, 맹검법(blinding)을 사용한다.

일곱째, 측정기구를 표준기(gold standard)를 사용하여 주기적으로 보정한다.

5-2. 진단검사의 타당성(Validity of Diagnostic Tests)

1) 감수성(Sensitivity)과 특이성(Specificity)

; 진단의 타당성이란 질병이 있는 사람과 없는 사람을 구분하는 능력이라고 정의할 수 있다. 진단이란 결국 질병이 있느냐 또는 없느냐 하는 것을 판정하는 일인데, 진단검사법의 결과는 대개 세 가지 척도(명목척도, 서열척도, 간격척도)로 표현된다. 따라서 진단검사법의 결과가 어떤 척도로 표현되든지 이를 2분변수(양성/음성, 정상/비정상)로 판정하여야 한다. 그러나 이러한 판정은 항상 옳을 수가 없으며, 어떤 진단검사법을 사용하든지 분류편의가 개재될 가능성은 항상 존재한다.

- 감수성: 질병이 있는 사람을 질병이 있다고 판정하는 확률
- 특이성: 질병이 없는 사람을 질병이 없다고 판정하는 확률

표 . 진단검사 결과

		질병 있음	질병 없음	계	
진단검사법	양성	a	b	a+b	$P(\text{유병률}) = (a+c)/N$
	음성	c	d	c+d	$P^*(\text{양성률}) = (a+b)/N$
	계	a+c	b+d	N	
특이성 = $\frac{d}{b+d}$		감수성 = $\frac{a}{a+c}$			

5-3. 감수성과 특이성의 교역관계 (Trade-offs between Sensitivity and Specificity)

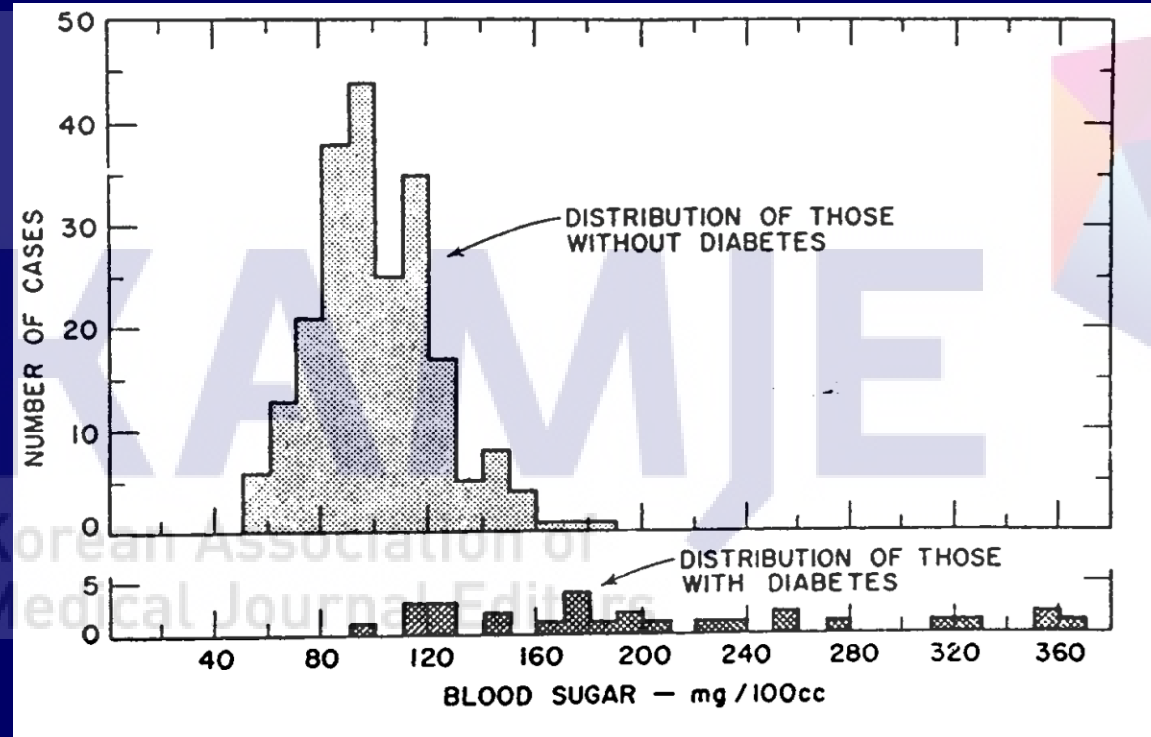


그림 . 당뇨병 환자와 정상인의 혈당치 분포

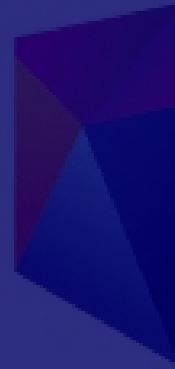
5-4. 감수성과 특이성에 영향을 주는 요인

1) 환자구성(Spectrum of Patients)

2) 편의(Bias)

3) 확률(Chance)

KAMJE
Korean Association of
Medical Journal Editors



5-5. 진단검사의 예측도(Predictive Value of a Test)

진단검사법	질병 있음		질병 없음		계	
	양성	a	b	a+b		
	음성	c	d	c+d		$P(\text{유병률})=(a+c)/N$
	계	a+c	b+d	N		$P^*(\text{양성률})=(a+b)/N$

$$\text{양성예측도} = a/(a+b)$$

$$\text{음성예측도} = d/(c+d)$$

1) 예측도에 영향을 주는 요인 (Determinants of Predictive Value)

$$P \times (\text{양성률}) = (a+b)/N = a/N + b/N$$

$$\begin{aligned} &= \left(\frac{a+c}{N} \times \frac{a}{a+c} \right) + \left(\frac{b+d}{N} \times \frac{b}{b+d} \right) \\ &= P \times S_n + (1-P)(1-S_p) \end{aligned}$$

S_n = 감수성
 S_p = 특이성
 P = 유병률

$$\text{양성 예측도} = \frac{a/(a+b)}{(a+b)/N} = \frac{a/N}{P \times S_n + (1-P)(1-S_p)}$$

TABLE 5-9 ▣ Calculations of Predictive Values for Neural Tube Defects (NTD)* for α -Fetoprotein (AFP) Test in High- and Low-Risk Women

	AFP Test	Pregnancy Outcome		Totals	Predictive Value (%)
		NTD	Normal		
High-risk women	Abnormal	87	18	105	82.9
	Normal	13	9,882	9,895	99.9
	Totals	100	9,900	10,000	
Low-risk women	Abnormal	128	179	307	41.7
	Normal	19	99,674	99,693	99.98
	Totals	147	99,853	100,000	

*Spina bifida or encephalocele.

From Sheffield LJ, Sackett DL, Goldsmith CH, et al: A clinical approach to the use of predictive values in the prenatal diagnosis of neural tube defects. Am J Obstet Gynecol 145:319-324, 1983.

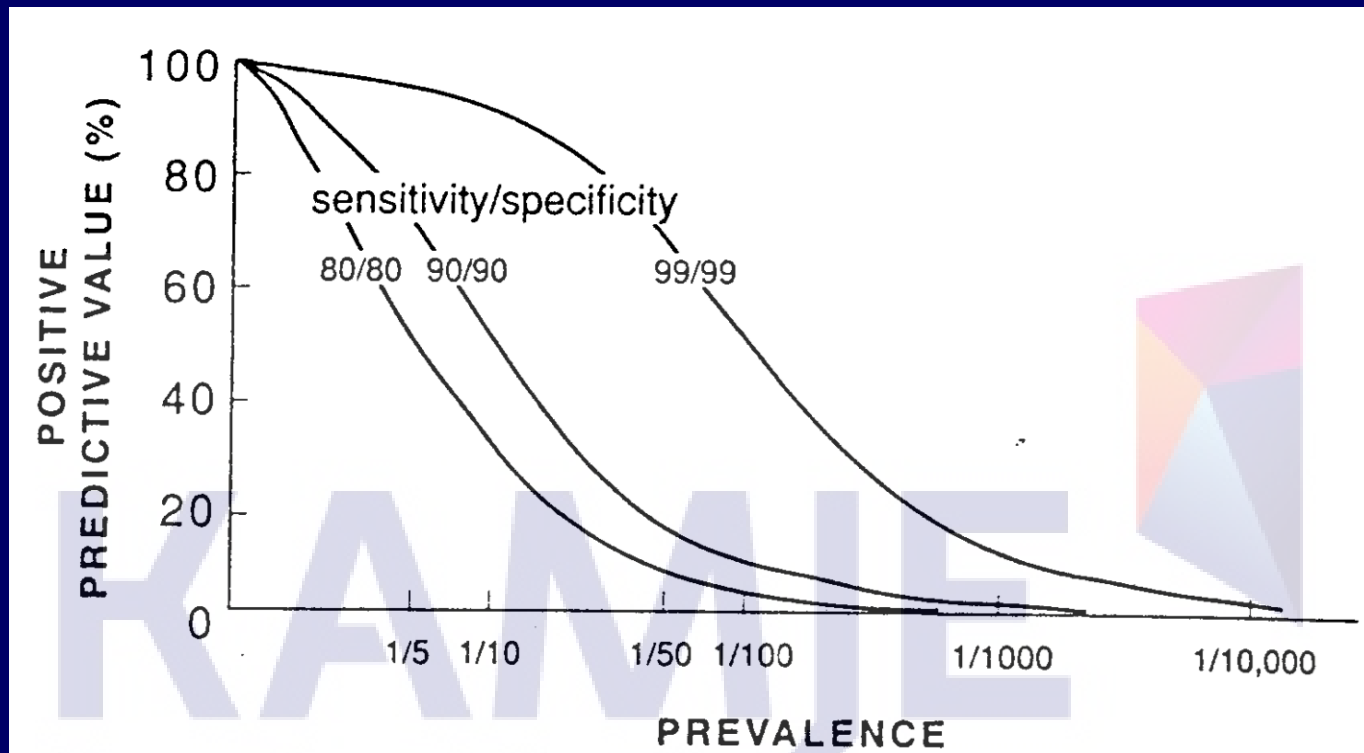
◦ 유병률과 양성 예측도와의 관계를 주목하여야 하는 이유

첫째, 선별검사의 효율성이 높아야 하기 때문

동일한 선별검사라 하더라도 유병률이 높은 집단
(고위험집단, high-risk group)에 적용하면 양성 예측도를 높일 수 있다.

**둘째, 검사 결과를 해석할 때 피검사자가 속한 집단의 유병률을
고려하여야 하므로**

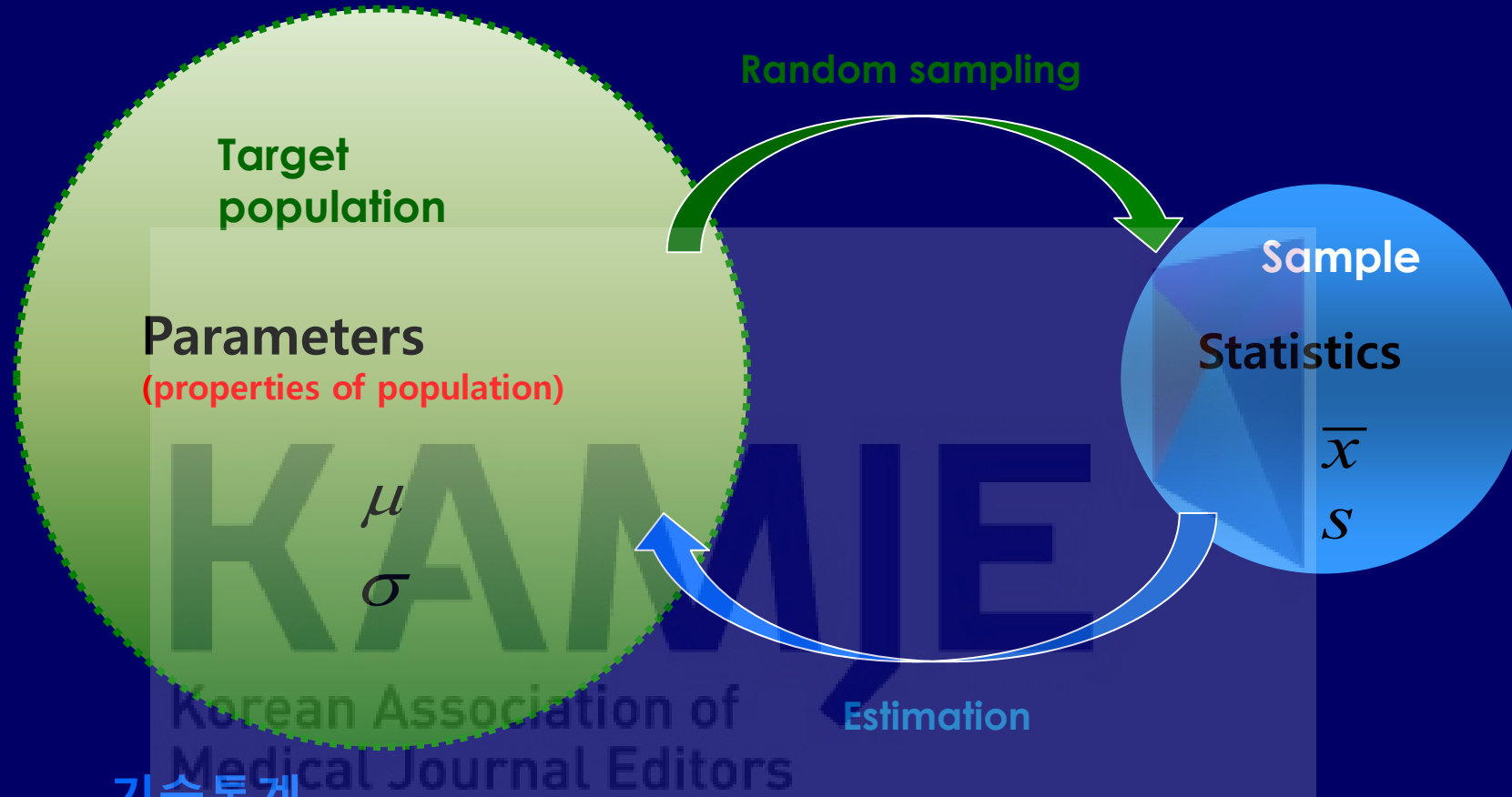
예) 양성으로 판정된 두 사람 중에 한 사람은 고위험집단에 속하고 다른
사람은 유병률이 낮은 집단에 속한다면 두 사람의 양성 예측도가 크게
차이가 날 것이므로, 동일한 양성이라도 그 해석을 달리 하여야
할 것이다.



Korean Association of
Medical Journal Editors

그림 14-5. 감수성, 특이성, 유병률과 양성 예측도의 관계

6. 분석방법



- 기술통계
- 추론통계
 - ✓ 점추정 (point estimation)
 - ✓ 구간추정 (interval estimation)

자료 성격	권고 통계분석 방법
사례보고, 임상연구, 치료결과분석 등	No statistical method or Descriptive study
진단능력평가, 참고치 정하기	Sensitivity, Specificity, ROC curve
짝을 이룬 두 그룹간 평균비교	Paired t-test, Wilcoxon signed rank test *
독립적인 두 그룹간 평균비교	t-test, Wilcoxon rank sum test *, Mann-Whitney U test *
독립적인 세 그룹 이상 평균비교 (또는 군간비교)	ANOVA (with multiple comparison), Kruskal-Wallis test *
동일인에 대한 3회 이상 반복측정자료의 평균비교	Repeated measures of ANOVA, Friedman test *
두 그룹 또는 세 그룹 이상 빈도 비교	Chi-squared test *, Fisher's exact test *
동일인에 대한 반복측정 빈도 비교	McNemar's test *
두 연속변수간 상관관계 분석	Pearson's correlation, Spearman's rho *
두 개 이상 독립변수와 종속변수와의 관계 분석	Simple linear regression, Multiple (logistic) regression
생존율 추정, 생존율 비교 생존형 자료의 회귀분석	Life table, Kaplan-Meier method Log-rank test, Cox's proportional hazard model (HR)
역학적 통계량 분석	Incidence, Prevalence, Risk ratio (RR), Odds ratio (OR)

* 비모수적 방법

Source: Emerson JD, Colditz GA. Use of Statistical Analysis in The New England Journal of Medicine. *N Engl J Med* 1983; 309: 709-713.

7. AI가 도와줄 수 있는 사항

- STROBE, CONSORT, PRISMA 등의 체크리스트 검토
- STROBE (관찰연구) ; 연구설계 명시 여부, 대상자 선정기준 기술 여부, 변수 정의 여부, 결측치 처리 방법, 통계분석 방법 등
- 다음 Methods를 STROBE 기준으로 평가해줘
; 누락 항목, 부족한 설명, 추가 기술 필요 부분 등을 빠르게 찾을 수 있다.
실제로 초고 단계에서 매우 유용하다.

8. 주요 메시지

□ AI는 Checklist reviewer, Writing assistant, Statistical assistant 가 될 수는 있다.

□ 연구자 고유 업무

; 연구질문을 만들고, 연구를 설계하고, 결과를 해석하고,
결론에 책임지는 것

결론; 독자는 연구방법을 왜 읽는가?

- 독자는 연구방법의 문장을 읽기위해 논문을 보는 것이 아니다.
- 독자는 연구결과를 믿어도 되는지 판단하기 위해 연구방법을 읽는다
- 좋은 연구방법의 조건
 - ; 연구 대상이 명확하고, 연구설계가 적합하고,
측정이 정확하고 구체적이며 신뢰도와 타당도가 확보해야 하며,
분석방법이 적절해야 한다.