

# The reviewer's comments (2): Statistical perspectives

안형진 교수  
고려대학교 의과대학

# Introduction

- 의학 학술지에서의 통계 리뷰
  - 많은 의학 학술지에서 통계 리뷰를 실시하고 있음.
    - Harwicke(2020)의 조사에 의하면 응답한 107개의 바이오메디컬 저널에서 66%는 전문 통계 리뷰를 하고 있고 이 중에서 23%는 모든 논문에 적용하고 있음.
  - 통계방법을 수행한 경우 통계 리뷰과정이 필요함 (적어도 두번째 리뷰 과정에서라도)
  - 통계리뷰 과정은 의학통계 전공의 전문가가 수행하는 것이 바람직함
  - 통계리뷰 과정에 대한 가이드라인을 학술지에 소개하는 것이 필요함 (체크리스트 포함)

# Common Statistical Comments

- 방법(method)부분에 연구설계와 자료수집과정을 자세하게 기술
- 분석한 방법의 명확하고 정확한 기술
  - “통계 분석”부분을 기술할 때 독자가 이 부분을 읽고 원자료가 있다면 논문의 결과를 다시 재생산할 정도로 자세하게 기술하는 것이 일반적 원칙임.
  - 통계분석방법은 방법의 이름만을 나열하는 것이 아니라 어떤 연구가설을 보이기 위하여 어떤 방법을 사용하였는지 자세히 기술
  - 올바른 통계분석법의 이름을 사용해야 함. (multivariate analysis for multiple regression, impaired t-test etc.)
  - 통계적 유의수준 지정

# Common Statistical Comments

- 방법(method)부분에 연구설계와 자료수집과정을 자세하게 기술
  - 결측자료 처리 방법
  - 변수의 선택 (모형 선택) 방법
  - 분석에 사용한 통계프로그램 및 버전의 명시
  - 임상시험과 관찰연구는 설계 방법이 다르기 때문에 그 설계에 맞는 방법을 구체적으로 기술하여야 함.
  - 만일 기계학습 방법이나 베이지안 방법 등의 익숙하지 않고 복잡한 방법의 경우는 관심있는 독자를 위하여 논문의 부록(appendix)에 좀 더 자세하게 기술함.

# Common Statistical Comments

- 설계 부분
  - RCT에서 무작위 배정의 오류
    - 월,수,금에는 처치군을 화,목,토에는 대조군을 배정하는 것은 무작위 배정이 아니고 체계적 배정임.
  - 적절하지 않은 대조군의 이용
  - 적절한 표본수 및 근거
  - 응답률의 적절성

# Common Statistical Comments

- 통계 분석

- 연구결과의 질은 얼마나 많은 통계분석방법을 사용하였느냐 또는 얼마나 어려운 통계분석방법을 사용하였느냐에 결정되는 것이 아니라 얼마나 적절한 방법을 사용하였느냐가 중요함.
- 반복측정 자료에서 상관을 고려하지 않은 분석 (예: independent two sample t-test for paired data)
- t-test를 이용한 생존시간 비교 (생존시간은 분포가 치우쳐 있고 중도절단이 있음)
- P-value의 적절하지 않은 해석

# Common Statistical Comments

- 통계 분석

- 연속형 변수의 범주화

- 중앙값이나 사분위수를 이용한 범주화는 자료에 따라 달라지기 때문에 좋은 방법이 아님.
    - 임상적으로 의미가 있어야 함.
    - 범주화 방법에 따라 분석 결과가 많이 달라질 수 있음.

- 관찰연구에서 교란변수(confounders)의 선택 및 보정방법

- 결과에 편향을 최소화할 수 있도록 교란변수를 선택하고 보정해야 함.
    - 측정되지 않은 교란변수와 잠재적 편향에 대해서는 Discussion에 기술되어야 함.
    - 성향점수매칭 → 좋은 방법이지만 매칭 이후에 충분한 검정력을 확보해야 함.

# Common Statistical Comments

- 통계 분석
  - 제 1종의 오류 통제
    - 다중비교
  - 적절하지 않은 사후 부집단 분석
    - 미리 계획되지 않은 사후 부집단 분석은 적절하지 않음.
    - 특히 유의한 결과를 찾을 때 까지 부집단으로 나누어 분석하는 것은 피해야 함.
  - 사용한 분석의 가정을 검토하지 않음.
    - 가장 흔한 오류이지만 통계 리뷰어가 검토하기 어려운 부분임.



# Common Statistical Comments

- 결측 자료의 처리
  - 결측 자료의 처리 방법을 구체적으로 기술함.
  - 결측 자료를 대체하는 방법을 사용한 경우는 대체한 구체적인 모형이나 방법을 제시. 단순히 결측자료 대체 방법을 제시하였다고 기술하는 것은 충분하지 않음.
  - 특별한 경우 (허가용 임상시험)를 제외하고는 Last Observation Carried Forward 방법은 사용하지 않는 것이 좋음.

# Common Statistical Comments

- 적절한 통계방법의 선택 기준
  - 연구의 목적 (가설)
  - 연구설계방법
  - 분석 변수의 수
  - 비교하고자 하는 군의 수
  - 자료의 종류 (연속형, 이분형, 범주형, 생존기간 등)

# Common Statistical Comments

- 결과 (논문에서의 표현) 부분
  - 자료의 분포적 특징
    - 기술통계 (그림 또는 요약숫자)
    - 연속형 자료: 평균±표준편차, 중앙값±사분위범위 (가능하면 둘 다 제시)
    - 범주형 자료: 절대빈도 및 상대빈도 (%)
    - 이상값 및 특이값 검출
    - 특이값의 제외 시 그 이유를 논문에 기술

# Common Statistical Comments

- 결과 (논문에서의 표현) 부분
  - 결과는 표에 제시된 숫자를 다시 반복하는 것이 아니라 가능하면 표에 제시된 값들의 질적인 표현에 중점을 둬
  - 자료의 분포에 대해 기술하는 경우(일반적으로 표1)는 SE가 아닌 SD를 제시함.
  - 원그림(pie-chart)는 논문에서 사용하지 않음.
  - 단지 p-value만 기술함. (적절하지 않음.)
  - 군간 비교의 경우 군별 추정치 및 CI뿐만 아니라 군간 차이 (또는 OR 등)의 추정치 및 CI를 제시함.
  - 신뢰구간을 제시하는 경우 (a, b) 또는 (a – b)보다는 (a to b)가 더 적절함.

# Common Statistical Comments

- 결과 (논문에서의 표현) 부분
  - 너무 적거나 또는 많은 소수점
  - 부적절한 P-value의 표현 ( $p=NS$ ,  $p=0.000$ , p-value 없는 \*, \*\* 등)
    - P-value는 가능하면 있는 그대로 표현 (소수점 둘째 자리 또는 셋째 자리까지, 예:  $p=0.10$ ,  $p=0.041$ ,  $p=0.001$  등)
    - 매우 작은 P-value는  $<0.001$  등으로 표현
  - 정확한 숫자 확인 (예: 표에서 총합이 맞지 않는 경우 등)
    - 어떤 표에서 숫자가 정확하지 않으면 다른 결과들도 의심하게 됨.
  - 그림으로 결과를 표현하는 경우에는 x축과 y축의 단위를 포함한 변수 설명을 명확히 제시하고 범례(legend)를 구체적으로 표시
  - 표에서 약어를 사용하는 경우 주석에 풀어서 제시
  - 그림과 표의 핵심은 표나 그림만으로 저자가 표현하고자 하는 바를 독자가 이해할 수 있도록 함.

# Common Statistical Comments

- 결과 해석 부분
  - 유의하지 않음을 차이가 없음이나 영향이 없음으로 해석하는 경우
  - 연관성을 강한 인과관계로 해석
  - 연구 자료 분석결과와 관련 없는 결론
  - 표본수가 작은 연구에서 추정치를 그대로 해석 (예: 처치군이 대조군에 비하여 질환에 걸릴 위험이 20배 → 생물학적으로 거의 불가능)
  - 표본수가 너무 큰 연구에서 P-값에 너무 의존한 해석 (예: 건강보험자료 분석 결과는 거의 대부분의 P-값이 0.05보다 작음. 임상적 유의성과 통계적 유의성은 항상 같은 것은 아님)

# 결론

- 의학연구를 수행하고 타당한 결과를 도출하여 논문으로 출판하기 위해서는 먼저 명확하고 의미 있는 연구주제를 확립하고 이 연구주제에 맞는 올바른 연구설계를 하여야 함.
- 연구설계대로 자료를 수집하고 적절한 통계분석법을 이용하여 결과를 내고 논문에 명확하게 통계방법과 결과를 기술함.
- 통계 리뷰는 논문의 질을 향상시키고 과학적 결론을 도출할 수 있도록 도움을 줄 수 있음.

# Reference

- Hardwicke and Goodman (2020), How often do leading biomedical journals use statistical experts to evaluate statistical methods? The results of a survey. PLoS One. 2020; 15(10): e0239598.