

이거나, 셋째, 경구 항알레르기제의 연간 외래처방일수 21일 이상 이거나, 넷째, 흡입 항염증제의 연간 외래처방일수 5일 이상인 경우였다. 이 기준에 따른 이차 조작적 정의는 연간 천식주상병으로 입원하였거나 종합병원 외래를 4일 이상 방문한 경우였으며 병원과 의원급에서는 연령군에 따라 구분기준이 달라졌다. 이상의 조작적 정의에 따라 산출된 연간(2002년 7월~2003년 6월) 천식환자 발생률은 0.601%, 유병률은 1.793%, 천식유병환자 사망률은 0.906%, 평균 유병기간은 2.8년이었다. 이 연구에서 산출한 유병률과 2003년도 국민건강영양조사, 2003년 건강보험통계연보에서 산출한 유병률들을 2004년 추계인구에 적용하여 유병환자수를 추계한 결과 각각 954,551명, 687,356명, 2,042,432명이었다.

결론: 이 연구에서 건강보험자료원을 이용한 천식환자수 추정방법은 과소 또는 과대평가 가능성이 높은 기존결과치들의 중간수준에 속하였지만 이 결과가 참값에 더 가깝다고 평가할 근거는 없다. 다만, 건강보험자료원의 구조적 제한점을 이해하고 일관된 조작적 정의를 적용한다면 장기간 추세, 일별-지역별 변이, 의료이용 과정 및 건강결과와 연계된 분석 등 기존 자료원에서 산출하기 어려웠던 활용도가 높은 정보들을 확보할 수 있다는 장점이 있으므로 향후 타당도를 검증하고 정확도를 높이기 위한 후속 연구가 필요할 것이다.



역학 6. 의료정보통계

KSPM-40

Statistical analysis of the schizophrenic behavior data with non-ignorable non-monotone missings

Sungcheol Yun¹⁾, Youngjo Lee²⁾, Moo-Song Lee³⁾, Sang-Il Lee³⁾, Young-Ho Khang³⁾, Min-Woo Jo³⁾

(1) Asan Medical Center, Office for Biostatistics Researches, (2) Department of Statistics, Seoul National University College of Natural Science, (3) Department of Preventive Medicine, University of Ulsan College of Medicine

Rubin and Wu (1997) presented the schizophrenic behavior data, which have many non-monotone missings with abrupt changes among repeated measures from a schizophrenic. Missing data are common in longitudinal studies and improper treatments of them can lead to wrong statistical inferences. Normal distribution can be assumed for the analysis of continuous responses. However, in biomedical researches outlying abrupt observations are far more frequent than suggested by the normal distribution. In this paper we introduce a simple interpretable model to account for non-monotone missingness with abrupt changes among repeated measures. Development of likelihood inferences for such models is hard because it involves intractable integration. We use hierarchical likelihood to overcome

such difficulty. Simulation study shows that the resulting estimator is efficient, robust against misspecification of fatness of tails and properly accounting for missingness. In the schizophrenic behavior data missingness is non-ignorable and abrupt changes among repeated measures from a schizophrenic can be well described by introducing random effects in the dispersion.

KSPM-54

SAS, R, MLwiN을 이용한 다수준 분석: 이분형 자료를 이용해서 (Multilevel analysis using SAS, R, and MLwiN in binary data)

김선영¹⁾, 김윤희¹⁾, 김옥진¹⁾, 서희경¹⁾, 김호²⁾

(1) 서울대학교 보건대학원

목적: 최근 국내의 연구에서는 사회적인 요인이 건강에 미치는 영향을 보기 위해서 다수준 분석을 많이 사용하고 있다. 다수준 분석은 개인변수와 집단변수를 한 모형 안에 포함하면서 집단간의 변이를 고려하기 때문에 집단의 효과를 올바르게 추정한다. 국내에서 다수준 분석을 사용하는 연구들이 증가하면서 다양한 분석도구를 활용하고자 하는 요구가 늘어가고 있는 추세이다. 그러나 흔히 사용하고 있는 SAS나 R을 이용한 분석방법은 잘 알려져 있지 않고, 이러한 경향은 종속변수가 연속변수가 아닌 이분형 변수인 경우에 더욱 심하다. 따라서 이 연구에서는 로지스틱 모형을 이용한 다수준 분석 방법을 도입하고, 모형에 따라 SAS, R, MLwiN을 이용한 분석과정과 결과를 제시하고 비교해 보고자 한다.

방법: 본 연구는 2001년 국민건강 영양조사 자료를 이용하여, 개인 소득을 보정한 상태에서 집단 소득이 흡연여부에 미치는 영향을 알아보고자 하였다. 다수준 분석에 사용된 모형은 랜덤절편모형(random intercept model)과 랜덤기울기모형(random slope model)이며, 소득을 개인수준과 집단수준을 모두 고려하여 맥락효과(contextual effect)를 볼 수 있는 모형을 첨가했다. 따라서 본 연구에 고려된 모형은 1)영 모형(null model), 2) 랜덤절편모형(random intercept model), 3) 집단변수를 첨가한 랜덤절편모형, 4) 랜덤기울기모형(random slope model), 5) 맥락효과를 고려한 랜덤기울기모형이다. 이 모형을 분석하기 위해 SAS의 NLmixed procedure, R의 glmmPQL function, MLwiN을 사용했다.

결과: SAS, R, MLwiN에서 5개 모형을 분석한 결과, 추정된 고정 효과와 랜덤효과들의 분산은 유사하였으며, 이용된 프로그램 간에 차이는 크지 않았다. 예를 들어 모형3의 경우 집단 소득 효과의 추정치는 -0.003(SAS), -0.003(R), -0.003(MLwiN)로 유의한 효과를 보였으며, 이에 대한 랜덤효과들의 분산은 0.463(SAS), 0.433(R), 0.452(MLwiN)로 나타났다. SAS NLmixed procedure의 경우 랜덤 효과를 두개 이상 고려하지 못하는 한계가 있어서 모형4와 모형5의 경우 프로그램 간의 결과 비교는 하지 못했다.

결론: 다수준분석을 위해 본 연구에서 사용한 세 가지 통계프로그램은 모형의 적용이나 분석과정에 약간의 차이가 있었다. SAS는 한 개의 랜덤효과만 고려할 수 있었고, R과 MLwiN은 여러 개를 동시에 고려할 수 있었다. SAS는 모형의 계수들의 초기값을 설정해주어