

이거나, 셋째, 경구 항알레르기제의 연간 외래처방일수 21일 이상이거나, 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

야 하는 번거로움이 있지만, R이나 MLwiN은 초기값을 설정할 필요가 없다. MLwiN은 명령문을 입력하지 않고 대화창의 메뉴바를 선택하여 사용해 편리한 장점이 있지만, 분석과정을 저장할 수 없는 불편함이 있었다.

KSPM-109

국내 일부 주요 의학회지의 연구방법론 및 통계분석방법 활용에 관한 비교 연구 - 1992년과 2002년

A Comparative Study for Research Designs and Statistical Methods used in Major Korean Medical Journals - 1992 and 2002

이유진¹⁾, 이준영¹⁾, 김해준¹⁾, 김순덕¹⁾

(1) 고려대학교 보건대학원

Purpose: This study is to compare research designs and statistical methods used in selected journals from the three fields of medicine, namely, the basic, the general, and the clinical fields, published in 1992 and 2002.

Methods: Three journals of Korean Journal of Preventive Medicine (KJPM), Yonsei Medical Journal(YMJ) and Korean Journal of Gastroenterology(KJG) from each of the above-mentioned fields were analyzed to figure out its significant differences of the research designs as well as statistical methods. Articles in the journals were classified by research designs using Feinstein(1985)'s criteria, by its general purposes, by its concrete study goals, by the components of study group, and by statistical methods employed. To verify statistical association regarding study designs and statistical methods used in the journals and in its publication years, the chi-squared test and Fisher's exact test were used.

Results: Regarding research designs, KJPM contributed much to case-effect analyses and risk factor analyses, and YMJ contributed to descriptive studies and pathological analyses while KJG to case-report and pathological analyses. For all journals, cross-sectional studies were the most commonly used design types. Comparing to 1992, the proportion of case-effect researches in 2002 increased from 20.9% to 28.6% on their general purposes; risk factor analyses increased from 11.8% to 28.3% on concrete study goals; and surveys of therapy from 11.8% to 21.3% on their study designs. For statistical methods used in articles, a significant improvement was found in a decade. In 2002, the proportion of articles which used statistical methods were 91.2%, comparing to 85.7% in 1992. The average number of statistical methods used were 2.64, 1.90 and 1.72 per article in KJPM, YMJ and KJG, respectively. The most commonly used statistical method were the t-test and contingency tables regardless of journal types. In KJPM, however, more diverse and more advanced statistical methods were used than the others. Between 1992 and 2002, the proportion of articles which used advanced statistical methods has increased from 9.1% to 13.6%.

Conclusion: Noting that research designs in medical journals depend on their principles and characteristics of its own fields, there is a limitation of generalizing our study results to all medical journals. However, it can be seen that the statistical methods used in Korean medical journals have

been steadily improved over last 10 years.

KSPM-197

무선인터넷 연결 PDA를 활용한 임상시험 모니터링 방법에 관한 연구

A Study On the Clinical Trial Monitoring using Wireless Internet PDA

박숙경¹⁾, 박수현²⁾, 강민구³⁾, 채영문⁴⁾, 김석일⁵⁾

(1) 질병관리본부 예방접종관리과, (2) 가톨릭의대 내과학교실, (3) 한신대학교 정보통신학과, (4) 연세대학교 보건대학원, (5) 가톨릭의대 예방의학교실

목적: 무선인터넷은 인터넷의 즉시성과 휴대전화와 개인정보단말기(PDA)의 이동성을 합한 개념으로 그 활용범위가 점차 증가하고 있다. 제약업계에서는 임상시험의 총 업무부담을 감소시키고 프로세스 스피드를 증가시키는 등의 장점으로 웹을 통한 임상시험이 증가하고 있으며 우리나라에서도 인터넷을 활용한 임상시험 솔루션이 개발되어 진행되고 있다. 본 연구에서는 의료정보의 분야에서 무선인터넷을 임상시험에 활용한 새로운 시스템을 개발함으로써 무선인터넷이 쉽게 의료분야에 활용될 수 있다는 모형을 제시하고자 하였다.

방법: 본 연구는 서울지역에 위치한 약 800명상 규모의 대학병원에서 과민성 장증후군으로 외래를 방문한 환자 중 내복약 R의 임상시험 증례기록부의 피험자 선정기준 및 제외기준에 적합하고(부록), 임상시험에 참여하기로 동의한 환자 64명을 대상으로 하였다. 시험은 1주의 투약전 기초기간과 4주 동안 위약 혹은 내복약 R 1일 1정을 무작위 배정, 이중맹검으로 투여하는 기간으로 이루어진다. case에게는 PDA phone을 통한 e-CRF에, control은 증례기록부(서면)에 피험자가 판단한 지난 한 주 동안의 전반적인 증상 완화정도를 평가하여 입력한다. case는 PDA phone을 이용해 매일의 임상시험 유효성평가변수를 입력하면 모니터요원이 PDA phone이나 홈페이지를 이용해 데이터를 모니터링하고 관리한다. 입력된 데이터는 데이터베이스에서 실시간 관리정보를 분석하여 조회하였다.

결과: 초기 연구참여자는 모두 64명이었으나 중도탈락자를 제외하고, 최종연구대상자는 PDA 사용자 18명, 서면 사용자 18명으로 총 36명이었다. 피험자에게 매일 1회 유효성평가 변수를 기록하게 하고 그 응답률을 본 결과 전체 630건 중 서면을 이용한 군은 매일 입력한 건수가 595건(94.4%)이고, PDA를 이용한 군은 487건(77.3%)으로 서면을 이용한 군에서 응답률이 높게 나타났다($p=0.0001$). 그러나, 각 문항의 응답완성도를 보면 PDA를 이용한 군이 서면보다 응답의 충실도가 더 높게 나타났다($p=0.0001$). 입력오류의 경우, 서면이 전체 78건, PDA가 52건이었고, 1인당 입력오류건수는 서면이 7건, PDA가 4건으로 서면 사용자의 입력오류가 더 많은 것으로 나타났다. PDA 사용자에게는 약물복용후 다음날 오후2시까지 임상증상을 입력하도록 지시하였으며, 전체 입력건수 중 지시한 시간 내에 입력한 건수는 365(74.9%)였고, 지시한 시간외에 입력한 건수는 122(25.1%)로 나타났다. 서면 사용자의 경우 매일 임상증상을 입