

번호: OP-K-004					
제 목	당뇨관련 역학지표 산출에 필요한 표본추출을 위한 시물레이션 A Simulation Study for Determining Sampling Strategy to develop Epidemiological Indices Regarding Diabetes				
저 자 및 소 속	이지성1), 이준영2), 송주원3), 김재용4), 최인정4) 1) 고려대학교 대학원 보건학 협동과정 보건통계학 전공, 2) 이준영, 고려대학교 의과대학 예방의학교실, 3) 송주원, 고려대학교 정경대학 통계학과, 4) 건강보험심사평가원 조사연구실 Ji Sung Lee1), Juneyoung Lee2), Ju Won Song3), Jaiyong Kim4), In Jung Choi4) 1) Division of Biostatistics, Department of Public Health, Korea University, 2) Department of Preventive Medicine, College of Medicine, Korea University, 3) Department of Statistics, Coegee of Political Science and Economics, Korea University, 4) Health Insurance Review Agency				
분 야	역 학 [고혈압/당뇨]	발 표 자	이지성 일반회원	발 표 형 식	구 연
<p>목적: 본 연구는 2003년도에 1회 이상 주상병이나 부상병 혹은 기타 추가상병으로 건강보험심사평가원에 당뇨병이 청구된 경험이 있는 환자 250만명에 관한 자료를 대상으로 실시되었다. 먼저 검사, 시술 및 약제를 포함하여 당뇨병명이 붙은 모든 의료기관의 청구내역 및 외래를 통해 청구된 약 정보까지 결합하여 만든 환자 - 요양기관 프로파일인 3,676,164건의 데이터베이스를 구축하였다. 본 연구에서는 이 자료를 대상으로 당뇨 유병율을 산출하기에 적절한 표본설계 및 표본추출의 타당성을 검증하기 위하여 시물레이션이 실시되었다. 연구의 목적은 (유사)모집단을 알고 있는 상태에서 시물레이션을 통해 표본추출 방법 및 적정 표본 수를 평가하기 위한 것이다.</p> <p>방법: 환자 - 요양기관 프로파일 자료 3,676,164건 중 한 환자가 요양기관 1곳만을 방문한 경우는 1,768,600명(1,768,600건), 요양기관 두 곳 이상을 방문한 환자는 735,154명(1,907,564건)이다. 전국적으로 조사대상 기관 수는 종합전문병원 42개소, 종합병원 248개소, 병원 944개소, 그리고 의원 20,409개소였다. 표본추출 방법으로는, 먼저 전국을 일정 기준에 따라 층화한 뒤, 각 층 내에서 1단계로 의료기관을 먼저 선택하고, 선택된 의료기관 내에서 환자들을 확률추출하는 층화 2단 집락추출(stratified two-stage cluster sampling)을 실시하기로 하였다. 층화(stratification)의 방법으로는 ① 전국을 구/시/군 단위로 구분하는 방법 ② 16개 광역시도로 구분하는 방법 ③ 9개 도(내륙 8도 및 제주)로 구분하는 방법 등을 고려하였다. 이들 각 구분 방법에 따른 “종합전문병원 : 종합병원 : 병원 : 의원”의 비례배분은, 기관의 수, 외래일수 등등을 종합적으로 고려하여 ① 5:5:5:85 ② 10:10:10:70 ③ 5:10:10:75 등의 비례배분을 조사하였다. 요양기관 두 곳 이상을 방문한 환자들의 경우, 이들의 대표 요양기관을 선택하는 기준으로는 ① 환자 별 최다 외래이용일수를 보인 요양기관을 선택하는 방법 ② 혈당강하제 그리고/또는 인슐린 처방일수가 제일 긴 요양기관을 선택하는 방법 등을 고려하였다. 요양기관 선택 후, 의원급 요양기관 중, 당뇨 포함 입원 및 외래일수 중 당뇨관련 외래일수 합계의 비율이 1% 미만인 의원, 그리고 종별 당뇨환자 수가 40명 미만인 요양기관(의원은 30명 미만)은 분석 대상에서 제외하였다. 표본추출 대상 환자 수의 선택 기준으로는 현실적인 면까지 고려하여 ① 2,000명 ② 4,000명 ③ 6,000명 ④ 8,000명 등을 선정하였다. 시물레이션은 위에 언급한 각 조건들에 따라서 층화 2단 집락추출된 표본들을 1,000번 반복하여 당뇨 유병율을 추정하였다. 최종 분석대상 자료는 2,412,082명이며, 표본 추출 및 유병율의 계산은 SAS version 8.1 및 SUDAAN version 9.0을 이용하였다.</p> <p>결과: 환자 - 요양기관 프로파일 자료에 근거한 당뇨유병율의 모수(모수의 근사치)는, 건강보험심사평가원 조사연구실 및 대한당뇨학회 역학소위원회의 협의에 따라, 전체 환자들 중 당뇨병이 한번이라도 청구된 적이 있으면서 당뇨용제(인슐린 그리고/또는 혈당강하제)를 1일 이상 처방받은 환자들의 비율로 정의하였다. 이 정의에 근거한 모수(모수의 근사치)는 56.66%였다. 한편 시물레이션 분석 대상 자료를 근거로 계산된 당뇨유병율의 모수(대리모수)는 요양기관 선택기준 중, 환자 별 최다 외래이용일수를 보인 요양기관을 선택하는 경우가 54.47%, 당뇨용제 처방일수가 제일 긴 요양기관을 선택하는 경우가 56.74%였다. 시물레이션 결과, 층화의 기준으로 구/시/군 단위의 구분을 하는 경우, 종별 기관 수의 비례배분을 5:10:10:75로 하는 경우, 최다 당뇨용제 처방일수에 근거해 요양기관을 선택하는 경우, 그리고 4,000명의 표본수를 고려하는 경우에 모수 예측값이 57.04% (95% 신뢰구간, 53.97-60.11; 10% 절삭범위, 50.60-62.89)로 얻어졌으며, 이는 만족할 만한 수치로 판단되었다.</p> <p>결론: 모집단에 관한 자료가 없는 경우, 적정 표본 수의 계산은 제 1종 오류, 검정력, 유효크기 등에 근거한 검정통계량의 분포 또는 신뢰구간을 이용하여 예상하게 된다. 그러나 모집단에 관한 자료가 존재하는 경우, 효과적인 표본추출 방법의 선택 및 적정 표본 수 계산은, 이론적 근거에 따른 공식을 적용하는 것 대신에, 해당 자료에 근거한 시물레이션을 통해 보다 효과적으로 파악될 수 있다. 본 연구에서는 이러한 관점에서 당뇨관련 역학지표 산출에 필요한 표본추출 방법을 평가하고 적정 표본 수 산출에 관한 정보를 도출하였다.</p>					