

번호: PO-EP-023					
제 목	데이터마이닝을 이용한 당뇨병의 고위험군 예측모형 구축 Modeling to Predict High Risk Group in Diabetes Mellitus, Using Data-Mining Technologies				
저 자 및 소 속	박중순1), 김순덕2), 이제숙1) 1) 고려대학교 대학원 보건학협동과정, 2) 고려대학교 의과대학 예방의학교실 Jongsoon Park1), Soonduck Kim2), Jesuk Lee1) 1) Department of Public Health Graduate School, Korea University, 2) Department of Preventive Medicine, College of Medicine, Korea University				
분 야	역 학 [만성질환 역학 및 건강 위험요인]	발 표 자		발 표 형 식	포스터
<p><b>목적:</b> 최근 경제성장과 생활습관 및 식이습관의 서구화, 노령인구의 증가 등으로 인해 당뇨병이 증가하고 있다. 본 연구는 데이터마이닝 기법을 활용하여 W시 J지역의 대규모 데이터베이스로부터 당뇨병에 가장 많은 영향을 주는 최적화된 요인을 찾아내고, 다양한 패턴을 분석하여, 당뇨병의 고위험군 예측모형을 제시하고자 하였다.</p> <p><b>방법:</b> W시 J지역의 2000년-2001년 국민건강보험공단 자료를 이용하였다. 분석 대상은 2000년 건강진단을 받은 20세 이상인 28,381명 중에서 건강진단시 당뇨병의 과거력이 있거나, 당뇨병 환자인거나, 공복시 혈당수치가 110mg/dl 이상인 자를 제외한 24,123명을 대상으로 한 후 동일한 사람의 의료보험 청구자료를 이용하여 2002년까지 당뇨진료를 1회이상 받은 사람을 당뇨병으로 정의하여 분석하였다. 분석방법은 Chi-square test와 데이터마이닝 분석기법 중 로지스틱 회귀분석과 의사결정나무분석을 하여, 당뇨병의 고위험군 예측모형을 제시하였다. 의사결정나무분석시 독립변수는 2000년의 건강진단 시 연령, 성별, 최종 의료보험료, BMI, 최고혈압, 최저혈압, 총콜레스테롤, <math>\gamma</math>-GTP로 하였고, 종속변수는 2002년도의 당뇨병 진단여부이었다.</p> <p><b>결과:</b> 당뇨병의 2년간의 발생률은 2.28%로 나타났으며, 교호작용을 고려한 로지스틱 회귀분석결과 당뇨병의 고위험군 예측변수는 연령이 증가할수록, BMI가 증가할수록, <math>\gamma</math>-GTP이상으로 나타났으며, 50대에서는 총콜레스테롤 199mg/dl 이하보다 총콜레스테롤이 200-239mg/dl인 경우에 당뇨병의 위험이 3.3배 증가하는 것으로 나타났다. 특히 연령의 경우 당뇨병의 위험이 20-30대에서보다 40대에서 2.45배, 50대에서 2.46배, 60대에서는 10.3배 증가하여 다른 요인보다도 연령이 당뇨병에 가장 큰 영향을 미치는 것으로 나타났다. 연령별 의사결정나무 분석결과 20-30대에서 당뇨병의 고위험군 예측변수는 총콜레스테롤, 최저혈압, BMI이었으며, 40대에서는 BMI와 성별, 50대에서는 최고혈압과 BMI, 60대에서는 최고혈압, <math>\gamma</math>-GTP, 총콜레스테롤, 성별로 나타났다.</p> <p><b>결론:</b> 비만이나 총콜레스테롤이상군, <math>\gamma</math>-GTP의 이상군 및 고혈압을 가진 고위험군들이 당뇨병으로 발전하지 않도록 이들에게 필요한 영양, 운동, 절주 등의 보건교육을 실시할 수 있는 기반이 마련되어야 할 것이며, 이를 통해 보다 효과적인 건강증진사업이 이루어져야 하겠다.</p>					