

데이터마이닝 기법을 이용한 퇴원환자 이탈연구

함승우*, 김은엽**

*한국원자력의학원, **고려대학교 보건과학대학
hsw@kcch.re.kr, key0227@korea.ac.kr

A study on the Hospital dropout of Discharged patients Using Data Mining

*S. W. Ham, **E. Y. Kim

*Department of Medical Record, Korea Institute of Radiological
& Medical Sciences

**Department of Health Administration, Korea Univ.

요 약

Introduction : 최근 의료기관 내 진료비 정산으로 인한 전산화사업이 적극 추진되면서 각 의료기관은 생성된 자료의 목적과는 무관하게 대량의 자료를 지니고 있으면서도 병원경영과 관련하여 자료를 활용하고 있지 못하다. **Objectives** : 병원 입원환자의 데이터를 분석하여 입원환자의 이탈요인을 파악하여 이탈을 방지하고 지속적인 의료서비스를 제공할 수 있도록 하고자 한다. **Method** : 본 연구는 일개병원을 퇴원환자를 대상으로 2006년 1월 1일~2007년 12월 31일 동안의 퇴원환자 데이터를 기반으로 연구되었다. 퇴원환자 38,359명의 병원데이터 베이스를 기반으로 동일 연구기간 외래 누적데이터 375,659건의 외래환자 데이터와 비교하여 퇴원환자 중 재입원 예약 및 외래진료 예약 부도여부를 비교하는 후향적 연구를 수행하였다. 2008년 1월 1일 기준 재입원 예약 및 외래진료 예약부도 발생여부를 기준으로 계속진료환자 그룹과 이탈환자그룹을 생성하였다. 중복을 제거한 총 대상은 3,503명이었다. **Discussion** :본 연구에서 얻어진 결과를 기초로 퇴원환자의 이탈률이 높은 환자그룹을 집중 관리하여 이탈을 방지하고 관심을 갖고 향후 잠재적 이탈 고객을 다시 유입할 수 있는 많은 방안들이 마련되어야 할 것이다.

1. 서론

1.1. Background

최근 의료기관 내 진료비 정산으로 인한 전산화사업이 적극 추진되면서 각 의료기관은 생성된 자료의 목적과는 무관하게 대량의 자료를 지니고 있으면서도 병원경영과 관련하여 자료를 활용하고 있지 못하다. 병원 데이터베이스 정보를 활용하여 환자 이탈 방지에 영향을 미치는 요인을 찾아 표준화된 데이터를 제공함으로써 업무 프로세스 개선하도록 하여 고객에게 질적인 의료서비스를 제공하고 있다.

최근 병원에서는 고객서비스를 증대하기 위하여 데이터 마이닝 기법을 사용하고 있다. 데이터 마이닝 기법이란 대량의 가공하지 않은 데이터로부터 알려

지지 않은 새로운 정보나 지식을 찾아내는 일련의 과정으로 최근에 병원경영 입장에서 데이터마이닝 기법을 이용한 고객서비스 향상에 활용하고 있는 실정이다.

1.2. Objectives

병원 입원환자의 데이터를 분석하여 입원환자의 이탈요인을 파악하여 이탈을 방지하고 지속적인 의료서비스를 제공할 수 있도록 하고자 한다.

2. Methods

본 연구는 일개병원을 퇴원환자를 대상으로 2006년 1월 1일~2007년 12월 31일 동안의 퇴원환자

데이터를 기반으로 연구되었다. 퇴원환자 38,359명의 병원데이터 베이스를 기반으로 동일 연구기간 외래누적데이터 375,659건의 외래환자 데이터와 비교하여 퇴원환자 중 재입원 예약 및 외래진료 예약 부도여부를 비교하는 후향적 연구를 수행하였다. 2008년 1월 1일 기준 재입원 예약 및 외래진료 예약부도 발생여부를 기준으로 계속진료환자 그룹과 이탈환자그룹을 생성하였다. 중복을 제거한 총 대상은 3,503명이었다.

퇴원환자의 일반적 특성을 분석한 이후 로지스틱 회귀분석, 신경망모형, 의사결정 모형의 데이터마이닝 분석을 실시하였다. SPSS ver 15.0 & SPSS clementine ver 10.0을 사용하였다.

3. 결과

1.1. 일반적 특성, 내원상태, 진료상태에 따른 이탈환자 분석

성별에 따라 계속진료환자그룹은 남자 743명(55.7%) 여자 1,375명(63.4%)이었고, 이탈환자는 남자 592명(44.3%) 여자 793명(36.6%)로 나타났다. 이탈환자그룹 중 20대가 48명 42.9%로 다른 연령대와 비교하여 높은 이탈률을 나타냈다($p < 0.003$). 보험유형에 따른 이탈유무를 조사한 결과 산재보험으로 입원한 환자에서 90명(48.6%)이 가장 높은 이탈률을 보였다($p < 0.003$). 입원경로에 따른 이탈률을 조사한 결과 응급실로 들어온 환자가 외래로 입원한 환자의 이탈률 보다 높게 나타났다($p < 0.001$). 입원경험 유무에 따른 분석결과 처음 입원한 환자의 이탈이 다입원 환자보다 높게 나타났다($p < 0.002$). 수술유무에 따른 분석결과 수술하지 않은 환자가 수술하고 이탈한 환자보다 높게 나타났다($p < 0.018$). 암환자를 14암 질환군으로 분류 하였을 때 눈,뇌 및 중추신경계군은 이탈이 가장 높게 나타났다($p < 0.000$).

1.2. 입원환자의 특성

재원일수는 두 그룹모두 1-5일, 6-10일, 순이었으나, 특이하게 21일 이상 입원한 비율도 높았다($p < 0.001$). 퇴원과는 두 그룹모두 외과계가 내과계보다 높은 비율을 나타냈다($p < 0.020$). 환자유형은 일반환자가 높은 비율을 나타냈다($p < 0.000$). 진료예약은 예약보다는 미예약에서 높게 나타났다($p < 0.000$).

1.3. 퇴원환자 이탈 예측모형

1.3.1. 로지스틱 회귀분석을 통한 예측모형

퇴원환자의 요인 중 이탈여부에 유의한 영향을 미치는 독립변수는 성별, 연령, 지역, 수술여부, 진료예약, 환자유형, 퇴원과, 퇴원년 변수였다. 이탈환자 예측 정확도는 69.5%였다.

$$\cdot \text{Logit(이탈)} = -.49 - .14(\text{성별}) - .11(\text{연령}) + .17(\text{지역}) + .13(\text{수술여부}) - .45(\text{진료예약}) - .46(\text{환자유형}) - .27(\text{퇴원과}) - .43(\text{퇴원년})$$

1.3.2. 관별분석을 통한 예측모형

퇴원환자의 요인 중 이탈여부에 유의한 영향을 미치는 독립변수는 성별, 지역, 특진, 진료시간, 요일, 초재진여부 변수였다. 이탈환자 예측 정확도는 67.3%였다.

$$\cdot \text{진료함수} = -.75 - .05(\text{성}) - .04(\text{연}) - .06(\text{지}) - .01(\text{보}) + .03(\text{초}) - 0.02(\text{입}) + 0.02(\text{재}) - .05(\text{수}) - .18(\text{예}) - .02(\text{비}) - .18(\text{환}) + .09(\text{과}) + .17(\text{년})$$

$$\cdot \text{이탈함수} = -.83 - .08(\text{성}) + .06(\text{연}) + .10(\text{지}) - .02(\text{보}) - .04(\text{초}) + 0.04(\text{입}) + 0.03(\text{재}) + .08(\text{수}) + .28(\text{예}) + .03(\text{비}) + .28(\text{환}) - .14(\text{과}) - .27(\text{년})$$

1.3.3. K-평균 군집분석을 통한 이탈유형

환자유형에 따라 3군집으로 분류하였을 때 평균재원일수와 평균 진료비는 낮고 수술치료를 하지 않은 2군집에서 가장 높은 이탈을 보였다($p < 0.000$).

4. Discussion

본 연구에서 이탈환자에 대한 영향을 미치는 독립변수들을 이용한 환자 의료정보와의 연계를 통하여 이탈률을 최소화 할 수 있는 요소를 찾으면서 고객의 의료만족을 높일 수 있는 요소들 또한 찾아야 할 것으로 사료된다.

본 연구에서 얻어진 결과를 기초로 퇴원환자의 이탈률이 높은 환자그룹을 집중관리하여 이탈을 방지하고 관심을 갖아 향후 잠재적 이탈 고객을 다시 유입할 수 있는 많은 방안들이 마련되어야 할 것이다.

참고문헌

[1] 김광환, “데이터마이닝을 이용한 이탈환자 방지 모형에 관한 연구” 대한의료정보학회, 8(1), 221-223, 2002.
 [2] 김옥남, 강성홍. “데이터마이닝을 활용한 진료데이터 분석” 대한의료정보학회, 15(1), 194, 1999.
 [3] 알렉스 버슨, 스테판 스미스, 커트 티어링. “CRM을 위한 데이터마이닝” 대청, 2000.