

데이터 마이닝을 활용한 UHDDS의 표준간호용어 기반 기록의 분석*

Analysis of Uniform Hospital Discharge Data Set Using Data Mining Technique

Myonghwa Park, RN., PhD.*, Seok-Jo Yang, PhD,[†]

^{*}College of Nursing, Keimyung University,

[†] College of Engineering, Chungnam National University,
194 Dongsan-dong, Daegu 700-712, mhpark1@kmu.ac.kr

Abstract: The purposes of this study were to discover nursing specific knowledge with hospital discharge data set applying data mining technique and to identify the utilization of data mining skill for clinical decision making. Data mining based on enterprise miner was conducted on a large clinical data set containing standardized nursing language including NANDA, NIC, and NOC. Randomized 700 patient data were selected from year 1998 database which had at least one of the five most frequently used nursing diagnoses. Patient characteristics and care service characteristics including nursing diagnoses, interventions and outcomes were analyzed to derive the meaningful decision rules. Decision Tree Model using five most frequent nursing diagnoses as target value showed that nursing interventions and outcomes and other patient characteristics were classified reasonably and consistent to the nursing care plan suggested by experts. This study demonstrated the utilization of the data mining method through large data set with standardized language format to identify the contribution of nursing care to patient's health.

Keyword: Uniform Hospital Discharge Data Set, Data Mining, Standardized Nursing Language

I. 서론

최근 정보시스템의 급격한 발전과 더불어 대량의 데이터를 포함하고 있는 정보시스템을 구축 관리할 수 있는 데이터 웨어하우스가 병원정보시스템의 중요 요소로 자리잡고 있다. 이러한 데이터 웨어하우스 등을 통해 대용량 데이터의 저장, 관리가 가능하면서 부터 방대한 양의 데이터로부터 의미있는 정보를 발견할 수 있는 지식 획득 방법인 데이터 마이닝(data mining)은 의료정보를 효율적으로 활용할 수 있는 중요한 도구로 대두되고 있다¹⁾.

환자 데이터의 중요한 기본 자료인 퇴원요약정보는 환자의 퇴원기록을 대상으로 항목별로 입력한 자료로서 미국의 경우 국가차원에서 Uniform Hospital Discharge Data Set(UHDDS)을 만들어 관리하면서 환자 의료정보 데이터베이스의 가장 기본적인 항목을 설정하여 관리하고 있다²⁾.

전자간호기록의 발전과 간호용어의 표준화와 더불어 대용량 간호 자료의 축적 또한 가능해졌으며 UHDDS의 필수요소로 포함되고 있다. 대용량의 간호 자료의 데이터 마이닝을 통한 분석은 실제적인 환자 간호의 현상을 이해할 수 있게 하며 의미있는 간호정보와 간호지식을 획득할 수 있는 유용한 방법으로 활용할 수 있다³⁾.

본 연구는 UHDDS에 포함되어 있는 간호진단, 간호중재, 간호결과와 환자자료를 데이터 마이닝 기술을 활용하여 분석함으로써 대용량 간호자료에서의 간호지식 도출에 있어 데이터 마이닝의 활용가능성을 확인하고자 한다.

본 연구의 구체적인 목표는 다음과 같다.

- 1) 대용량 데이터베이스의 데이터 마이닝 과정을 확인한다.
- 2) 데이터 마이닝을 통해 간호진단, 간호중재, 간호결과와 패턴을 확인한다.
- 3) 간호패턴과 환자결과의 연관성을 확인한다.
- 4) 간호지식 도출에 있어 데이터 마이닝의 유용성을 확인한다.

II. 방법

1. 연구대상 UHDDS

연구대상 병원은 미국 중서부지역의 500침상 규모의 종합병원으로서 간호정보시스템을 통한 NANDA, NIC, NOC 자료가 퇴원요약정보에 포함되도록 구성되어 있다. 본 연구는 1998년 1월 1일에서 1998년 12월 30일까지의 성인환자의 퇴원요약정보를 분석의 대상으로 하였다.

2. 데이터 웨어하우스의 구성

본 연구에서 사용된 데이터 웨어하우스의 구성은 다음과 같다.

* 본 연구는 보건복지부 보건의료기술진흥사업의 지원에 의하여 이루어진 것임(과제고유번호: A050909)

Billing Number	Medical Service
Medical Record Number	Discharge Status
Social Security Number	Medical Diagnosis
Birth Date, Sex, Race	Nursing Diagnosis
Patient Zip Code	Nursing Intervention
Religion, Marital Status	Nursing Outcomes
Length of Stay	

3. 자료 분석 절차

본 연구의 자료분석은 다음과 같은 절차를 거쳐 수행되었다(그림 1).

- 1) 사용빈도가 높은 5가지 간호진단 선정: 1년간의 퇴원환자정보 중 가장 빈도가 높은 5개의 간호진단군(cluster)을 선정하였으며 선정된 간호진단군은 통증, 감염, 손상, 부동, 피부문제로 분류되었다.
- 2) 분석 대상 표본의 선정: 1998년 데이터베이스에서 선정된 5가지 간호진단군을 적어도 한 가지 이상 포함하고 있는 700개의 자료가 무작위로 선정되었다.
- 3) 간호중재의 코드화: 해당 중재가 기록된 경우를 1, 그렇지 않은 경우를 0으로 코딩하였다.
- 4) 간호결과의 코드화: 해당 결과가 기록된 경우를 1, 그렇지 않은 경우를 0으로 코딩하였다.
- 5) 자료 정리 및 처리: 중복 및 누락 자료를 정리하고 항목의 특성에 따른 분류와 재코딩을 실시하였다. 최종 데이터베이스는 700개의 열(row)와 66개의 attribute으로 구성되었다.

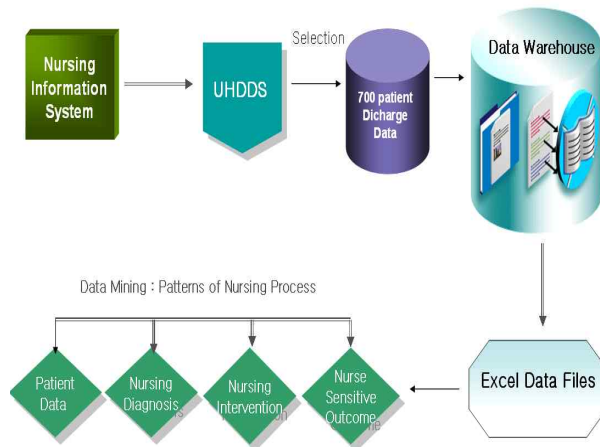


그림 1. 데이터 마이닝의 지식 획득 과정

3. 자료 분석 도구

대상 퇴원요약정보는 Access 파일로 전환되었으며 데이터 분석은 통계 패키지인 SAS와 데이터 마이닝 패키지인 Enterprise Miner를 이용하여 분석하였다. 분석에 사용된 기법은 의사결정 나무분석을 사용하였다.

의사결정 나무분석은 의사결정규칙(decision rule)을 나무구조로 도식화하여 분류와 예측을

수행하는 분석방법으로 예측의 과정이 트리구조에 의한 추론규칙에 의해 표현된다. 환자의 퇴원정보 특성과 간호특성 중 의미있는 지식을 발견할 수 있는 자료를 찾기 위해 5개의 간호진단군을 중심으로 관련되는 속성을 분류하는 과정을 거쳤다.

III. 연구결과

1. 연구대상의 일반적 특성

대상환자의 연령범위는 25-102세 사이의 분포를 보였으며 평균 연령은 70.31세, 재원기간의 평균은 5일로 나타났다. 환자들의 동반질환의 수는 평균 3.5개였다. 그 외 일반적인 특성은 표 1과 같다.

표 1. 대상환자의 일반적인 특성 (N = 700)

Variables	Category	N	%
성별	남	271	38.7
	여	429	61.3
퇴원분포	가정	430	61.6
	가정간호서비스	80	11.6
	장기요양시설	53	7.6
	타 병원으로 의뢰	10	1.4
	사망	122	17.6
입원병동	내과	251	35.9
	외과	218	31.1
	산부인과	118	16.9
	종양	36	5.1
	기타	77	11.0
DRG	순환기계질환군	209	29.9
	위장관계질환군	76	10.9
	정형외과군	92	13.1
	산부인과군	118	16.9
	기타	105	15.2

2. 간호진단, 중재, 결과의 분포

5개 간호진단군의 빈도는 표 2와 같다. 가장 빈번하게 사용되는 간호진단군의 순위는 감염, 통증, 손상, 피부문제, 기동성 장애의 순으로 나타났다.

간호중재와 간호결과 중 가장 빈번히 기록된 6개는 표 3과 표 4와 같다.

표 2. 5개의 가장 빈번히 사용된 간호진단군의 빈도 (N = 700)

간호진단군	N	%
Infection	593	84.7
Pain	504	72.0
Injury	202	28.6
Skin impairment	159	22.7
Immobility	100	14.3

표 3. 간호중재의 빈도 예 (N = 700)

순위	간호중재	N	%
1	Pain management	324	46.3
2	Infection protection	312	44.6
3	Analgesic administration	227	32.4
4	Skin surveillance	203	29.0
5	Medication management	109	15.6
6	Fall prevention	97	13.9

표 4. 간호결과의 빈도 예 (N = 700)

순위	간호결과	N	%
1	Remain free of infection	384	54.9
2	States realistic level of comfort	289	41.3
3	Maintain afebrile state	206	29.4
4	States relieved pain	195	27.9
5	Wound incision appear clean and free of purulent drainage	177	25.3
6	Identify intervention to prevent /reduce risk of infection	163	23.3

3. 간호진단별 의사결정 나무분석의 결과

간호진단 pain의 의사결정 나무분석(그림 2)은 간호중재인 pain management가 주요 중재로서 사용되며 주요 간호결과로서 "states realistic level of comfort"와 "states relieved pain"이 기록되었으며 환자의 입원병동과 DRG가 영향을 미치는 주요 속성으로 분류되었다.

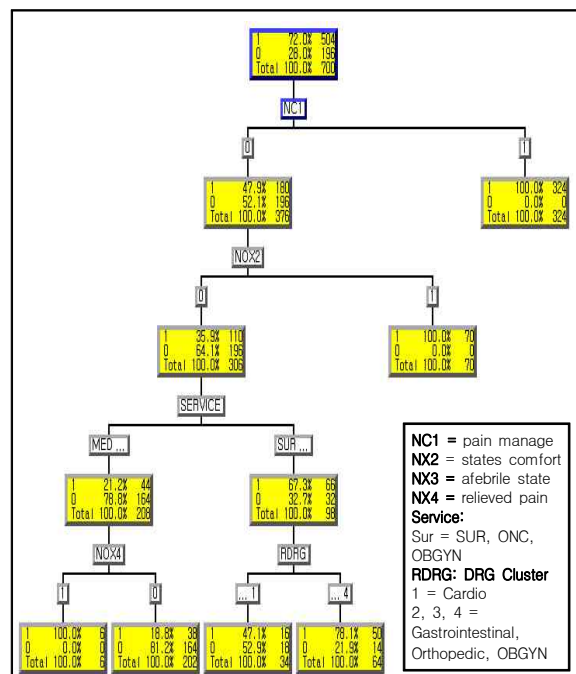


그림 2. pain 진단군의 의사결정 나무분석

간호진단 infection의 의사결정 나무분석(그림 3)은

간호중재인 infection protection이 주요 간호중재로 기록되었으며 주요 간호결과로서 "remain free of infection"와 "maintain afebrile", "intact dry skin", "wound incision appear clean"이 기록되었으며 환자의 DRG가 영향을 미치는 주요 속성으로 분류되었다.

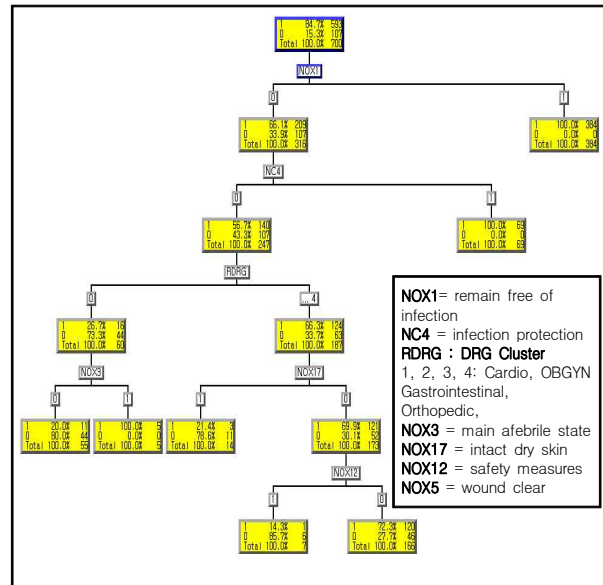


그림 3. Infection 진단군의 의사결정 나무 분석

간호진단 injury의 의사결정 나무분석(그림 4)은 간호진단 "infection"과 간호중재인 "fall prevention"과 "surveillance: safety"이 주요 속성으로 기록되었으며 주요 간호결과로서 "follow activity restriction"이 주요 간호결과로 기록되었으며 환자의 입원병동이 영향을 미치는 주요 속성으로 분류되었다.

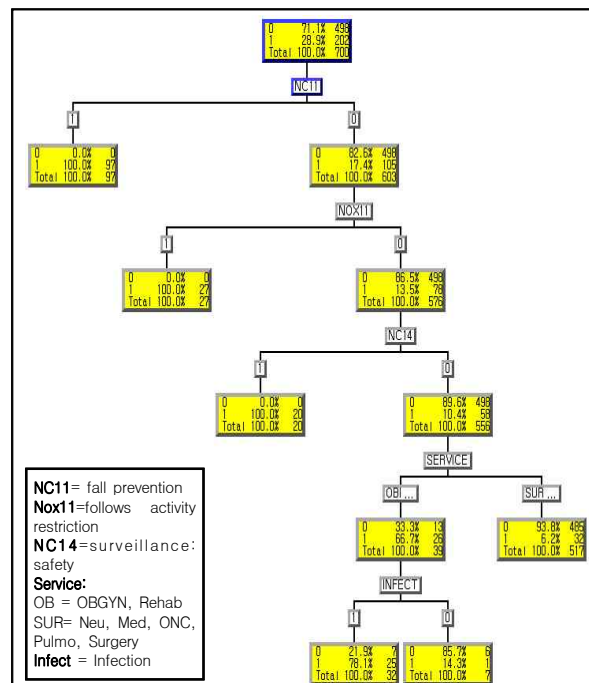


그림 4. Injury 진단군의 의사결정 나무분석

간호진단 immobility의 의사결정 나무분석(그림 5)

은 간호중재로서 "exercise therapy"와 "reposition", "environmental management"가 주요 속성으로 기록되었으며 주요 간호결과로서 "transfer with assistance/independently", "utilize safety measure" 등이 기록되었으며 환자의 DRG가 영향을 미치는 주요 속성으로 분류되었다.

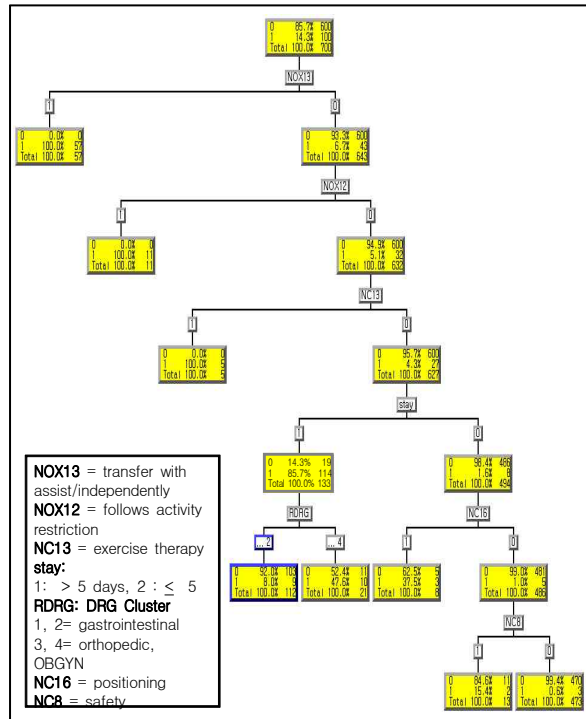


그림 5. Immobility 진단군의 의사결정 나무분석

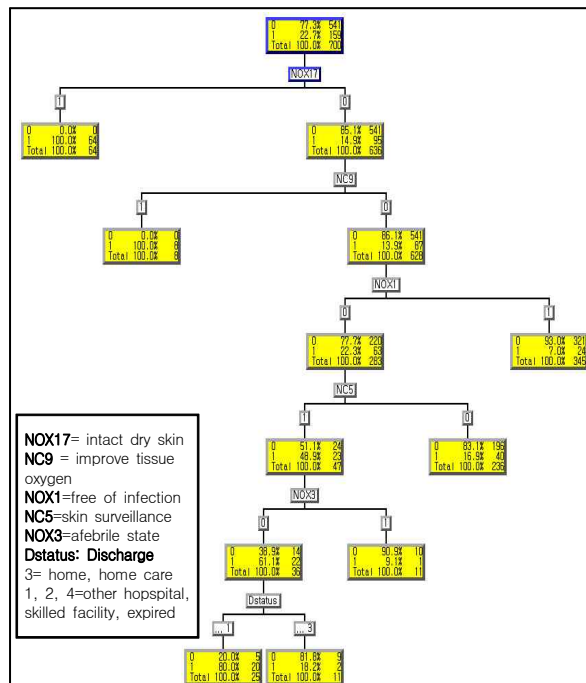


그림 6. Skin impairment진단의 의사결정 나무분석

간호진단군 skin impairment의 의사결정 나무분석 (그림 6)은 간호중재로서 "bedrest care"와 "skin surveillance"가 주요 속성으로 기록되었으며 주요 간

호결과로서 "intact dry skin", "remain free from infection", "maintain afebrile" 등이 기록되었으며 환자의 퇴원분포가 영향을 미치는 주요 속성으로 분류되었다.

IV. 결론

본 연구는 퇴원요약정보에 포함된 표준용어 기반 간호기록을 바탕으로 한 대용량 데이터를 데이터 마이닝 기법을 통하여 분석하여 의미있는 간호지식을 도출하는 연구의 일환으로 시도되었다.

본 연구의 결과는 간호사들이 실제로 수행 및 기록하는 간호과정의 요소들이 연구와 전문가 의견을 바탕으로 하여 문헌을 통해 제시된 간호계획과 상당히 일치하였으며 그 중 몇 가지 중요한 간호중재와 간호결과 및 환자 특성이 간호진단과 연관되어 있음을 확인하였다.

이러한 연구결과를 바탕으로 다음과 같은 제언을 하고자 한다.

첫째, 본 연구는 미국의 일개 병원을 대상으로 하였으나 국내 병원의 전산화와 더불어 대용량 간호정보가 구축된다면 국내의 자료를 바탕으로 데이터 마이닝을 적용하는 시도가 이루어져야 할 것이다.

둘째, 지속적인 추후연구를 통해 의미 있는 연관성 규칙들이 발견된다면 이를 바탕으로 한 critical pathway의 개발을 위한 연구가 이루어져야 할 것이다.

셋째, 환자의 질환명 혹은 DRG 그룹에 따른 간호진단, 간호중재, 결과의 규칙을 확인하는 연구가 이루어져야 할 것이다.

V. 참고문헌

- Goodwin L, Iannacchione MA. Data mining methods for improving birth outcomes prediction. Outcomes Manage, 2002;6(2): 80-85.
- Harris MR, Graves JR, Solbrig HR, Elkin PL, Chute CG. Embedded structures and representation of nursing knowledge. J Am Med Inform Assoc, 2000; 7(6): 539-549.
- Kraft MR. Mining a spinal cord injury clinical database for nursing information: A source of nursing knowledge. Unpublished Doctoral Dissertation, Loyola University of Chicago, 2002.