

데이터 마이닝 기법을 활용한 제왕절개수술환자의 Critical Pathway 개발에 관한 연구

호승희, 채영문¹, 최명애², 송미라³

한국과학기술원 테크노경영대학원, 연세대학교 보건대학원, 순천향 대학 병원 외무기학실, 메디메디아 코리아(주)

Study on the Development of Critical Pathway for the Cesarean Section using Data Mining Technique

Seung Hee Ho, Young Moon Chae¹, Myung Yae Choi², Mi Ra Song³

Graduate School of Management, Korea Advanced Institute of Science and Technology

Graduate School of Health Science and Management, Yonsei University¹

Soonchunhyang University Hospital, Medical Record Department²

Medimedia Korea³

Abstract

The objectives of this study was : to discover knowledge in predicting lengths of stay of Cesarean Section by patients' characteristics and treatment method using data mining technique : and to suggest the approach to the development of critical pathway. The findings suggest that data mining technique from the large pool of accumulated patients' data can be utilized to systematize newly observed correlations, patterns and trends and to develop critical pathway for the treatment and management. The results of the study can be contributed to aid developing the clinical pathway for cesarean section suitable to Korean patients. And the application of the developed critical pathway in clinical practice will produce the actual effect and value. (*Journal of Korean Society of Medical Informatics 8-2,41~50, 2002*)

Keyword : Cesarean Section, Critical Pathway, Lengths of Stay, Data Mining

1. 서론

최근 낮은 의료보험 수가와 포괄수가제의 적용으로 의료기관간의 경쟁 심화, 의료에 대한 국민의 권리 의식의 증가로 인하여, 의료비용의 절감과 질적 수준이 향상, 다른 병원과의 차별화 등 서로 상충된 목표를 달성하기 위하여 최근 Critical Pathway(CP)의 필요성이 대두되고 있다¹⁾. CP는 특정 진단명의 진료순서와 치료시점 등을 미리 정해 둔 표준화된 진료 과정을 일컫는다. 즉, 의사나 간호사 혹은 기타 의료진들이 어떤 진료 행위를 제공할 것이며, 그 진료 행위별로 어떤 시기에 제공할 것인지를 도식화한 것이다. 이러한 CP의 목적은 의료서비스의 지연과 자원 활용을 최소화하고 진료의 질을 극대화하는데 있다²⁾.

건강보험 심사평가원의 1999년도 통계연보에 의하면, 우리 나라의 경우 1999년 제왕절개분만 시행 건수는 196,808건으로 전체 분만 건수 중 43%를 차지하며, 비용은 66%에 달한다. 이러한 국내의 제왕절개분만율은 서구 국가와 비교시 3-4배가 높고, 85년 6.0%에서 98년도 상반기에 36.1%로 급증하였다³⁾. 그리고 2000년 국민건강보험공단 건강보험연구센터의 전국 진료비 청구자료를 분석한 결과에 의하면 제왕절개분만 시행 건수의 비율이 전체 분만 건수의 38.6%를 차지하고 있어 여전히 높은 비율을 나타내고 있음을 알 수 있다. 제왕절개의 경우 DRG를 시행하게 되어 있어 정해진 의료 수가 내에서 수술 및 처치를 해야 한다. 따라서 제왕절개 수술의 경우 비용절감과 의료서비스의 질적 수준 향상이라는 두 가지 목표를 동시에 얻을 수 있는 CP는 꼭 필요하다. CP는 주진료경로 과정상에서의 환자 진료 결과 목표를 설정하고 환자 진료 결과 목표의 달성 방법에 대한 가정을 설정하여 병원내 진료 현황 자료를 수집하여 치료의 분석을 통한 결과와의 관계를 규명하여 작성되어야 한다. 또한 임상진료지침 등 임상이론 및 해당 분야 임상 전문의의 견해에 근거해야 한다. 즉 개발시 충분한 근거를 가지고 최선의 치료방법을 도입하도록 해야 하며 이에 따라 환자의 특성을 반영한 다양한 pathway를 개발하여 치료가 너무 획일화되지 않도록 해야 한다⁴⁾. 이러한 과정이 지금까지와 같이 주관적으로 치우치거나 현재까지 알려

진 혹은 임상적 이론에 의거하는 것 보다는 방대한 데이터 웨어하우스로부터의 지식 획득을 통해 이루어져야 한다. 현재까지의 대부분의 CP 개발은 해당 분야의 임상진료지침, 전문가의 견해 등을 통합하여 의료진의 brain storming과 같은 합의과정을 거쳐 이루어짐으로써, 이의 개발 과정에 대한 현행 지식을 명확히 하고, 개발 과정의 변이 원인을 이해하고 개선 과정을 선정하는 데 있어 어려움이 있었다. 외국에서도 미국, 영국 등을 중심으로 1990년대 들어 evidence-based medicine이 새로운 패러다임으로 받아들여지면서 종래의 직관이나 비체계적 임상경험 또는 이론적(rational) 근거에 의존하던 의사결정 방식에 일대 변화가 예측되고 있다. 이에 따라 체계적인 합의 도출과정을 거치며, 광범위한 자료검색, 메타분석, 베이시안 분석, 비용효과분석 등의 계량적이고 명시적인 방법론을 활용한 임상지침 개발의 필요성이 대두되고 있다⁵⁾. 궁극적으로 의료분야에서도 다른 분야와 마찬가지로 의료자원의 효율적 활용을 위한 새로운 전략적 기법이 요구된다고 할 수 있다.

최근 정보시스템의 급격한 발전과 더불어 많은 양의 데이터가 축적되기 시작하면서, 데이터를 지식으로 효과적으로 저장, 관리, 활용할 수 있는 지식기반 경영시스템의 필요성이 대두되고 있다⁶⁾. 즉 단순히 데이터나 정보를 입력·조회·출력하는 역할이 아니라 지식을 바탕으로 지속적인 업무지식을 창출하고, 효과적으로 지식기반 경영시스템을 운용할 수 있는 전사적인 지식경영(knowledge management)의 비전 및 이의 전략수립이 요구되고 있다⁷⁾. 또한 최근 몇 년에 걸쳐 대량의 데이터를 포함하고 있는 정보시스템을 구축, 관리할 수 있는 데이터 웨어하우스(data warehouse)가 새로운 아키텍처로 각광받고 있으며, 이와 함께 방대한 양의 데이터로부터 함축적이며 잠재적 유용성이 있는 정보를 발견할 수 있는 지식탐사 방법인 데이터 마이닝(data mining)에 대한 중요도가 더욱 증대되고 있다⁸⁾. 데이터 마이닝은 패턴인식 기술이나 통계기법, 수학적 알고리즘을 이용하여 의미 있는 새로운 상관관계, 패턴, 추세 등을 발견하는 과정⁹⁾으로, 의료분야에서도 다른 분야와 마찬가지로 의료정보자원의 효율적 활용을 위한 지식기반 경영과 이의 요소정보기술로써 데이터 웨어하우스와 데이터 마이닝의 중요성이 부각되고 있다¹⁰⁾.

본 연구에서는 데이터 마이닝 기법을 활용한 진료 과정 및 결과 분석의 일환으로 제왕절개수술 환자의 데이터 마트를 구축하고, 데이터 마이닝 기법을 활용하여 재원기간을 결정짓는 규칙을 도출하고자 하였다. 그리고 도출된 규칙에 대한 결과를 분석하고 모형의 예측력을 평가하고자 하였으며, 향후 데이터 마이닝 기법을 활용한 CP 개발에 대한 방안을 제시하고자 하였다.

II. 연구 내용 및 방법

1. 연구 대상

서울 지역에 위치한 허기병상 약 1,500병상 규모의 대학병원에서 2000년도 4월에서 7월까지 퇴원한 의료보험 환자 중 주상병이 포괄수거제의 '제왕절개술' 질병군에 해당하는 261명을 대상으로 하였다.

2. 제왕절개수술 환자 데이터 마트 구축

해당 환자들의 퇴원요약 데이터베이스, 입원처방 데이터베이스, 임상병리검사 데이터베이스 등의 원본 데이터로부터 처방건수 기준으로 39,288건을 추출하고, 이를 변환, 통합하여 제왕절개수술 환자의 데이터 마트를 구축하였다. 이는 Access 데이터베이스로부터 요구 항목을 설정하여 MS-SQL을 사용하여 질의함으로써 이루어졌다. 그리고 제왕절개수술환자의 특성을 반영하는 변수들로 스키마를 구성하였는데, 이는 1개의 사실 테이블(fact table)과 적응증, 검사결과, 투약내역, 합병증, 치료결과 등의 5개의 차원 테이블(dimension table)로 이루어져 있다. 차원 테이블의 키에는 키 표가 있으며, 사실 테이블의 키는 차원테이블들의 키들로 구성되어 있다.

3. 데이터 마이닝 기법을 활용한 재원기간 결정 규칙 도출

CP개발 및 시행의 기대효과는 재원기간을 단축시키고 진료원가를 절감하면서도 환자의 만족도를 향상시키면서 양질의 진료서비스를 제공하기 위함이다. 이에 따라 사례관리에서의 적정성 평가 항목 및

CP의 재조정을 위해서는 목표 변수(target variable)를 '재원기간(lengths of stay), 진료비, 진료의 질', 세 가지 모두를 포괄할 수 있는 변수로 선정하여야 하지만, 본 연구에서는 이를 위한 일차적 접근으로 위의 세 가지 변수 중 재원기간을 목표 변수로 하여 이를 결정짓는 규칙을 도출하고자 하였다.

(1) 분석 모형

그림 1과 같이 데이터 마이닝 프로세스를 통해 환자의 특성과 처치 및 투약 내역, 합병증, 재원일수 간의 상관관계와 패턴을 발견하고자 하였다. 즉 재원기간의 분류(classification)와 예측(prediction)의 과정을 추론 규칙(induction rule)에 의해 표현하고, 이를 통해 대상군별 환자의 특성 및 처치 및 투약 내역의 특성을 분석하였다.

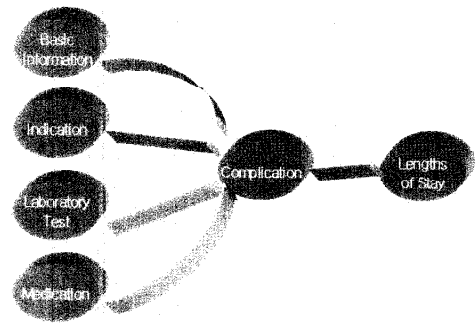


Fig 1. Framework of the analysis

(2) 분석항목의 선정 및 측정 수준

예측변수(predictor variable)의 설정을 위해 제왕절개술 환자의 특성 및 치료에서 중요한 특성으로 파악되어지는 항목들을 선별하였다. 이 과정은 임상진료지침(clinical practice guideline), 임상상의 견해, 실제 환자 데이터에 대한 기술통계량(descriptive statistics) 측정을 통한 데이터의 특성 파악 등을 통해 이루어졌다. 또한 선정된 변수는 측정 수준에 따라 연속형(continuous), 명목형(nominal)으로 분류하였다.

연령, 임신주수(intrauterine period), 신생아 체중(birth weight) 및 hematocrit은 연속형으로 측정하였으며, 제왕절개수술에 대한 적응증(indication)은 '반복 제왕절개술(repeat procedure), 난산으로 인한 분만

진행의 실패(dystocia), 둔위(breech presentation), 태아 곤란증(fetal distress)'으로 분류¹¹⁾하여 그 수준을 측정하였다. 그리고 투약내역은 항생제 투여 중 'penicillins계, cephalosporins계, quinolone계, macrolide계' 등으로 분류하였으며, 합병증은 '산욕기 감염(puerperal infection), 산욕기 발열(puerperal pyrexia), 산욕기중 정맥성 합병증(venous complication in the puerperium), 산과적 색전증(obstetric embolism), 산욕기중 마취제의 합병증(complication of anaesthesia during the puerperium), 제왕절개 부분 상처의 파열(disruption of caesarean section wound)'으로 분류하여 그 수준을 측정하였다. 또한 재원기간은 현재 Y병원에서 사용하고 있는 CP에서 제시한 6일을 기준으로 하여 6일 이하를 적정군, 6일 초과를 비적정군으로 분류하여 분석하였다.

(3) Decision tree를 이용한 재원기간 예측 및 대상군 분류

의사결정나무분석은 의사결정규칙(decision rule)을 나무구조로 도표화하여 분류(classification)와 예측(prediction)을 수행하는 분석방법으로 예측의 과정이 나무구조에 의한 추론규칙(induction rule)에 의해 표현된다¹²⁾. 본 연구에서는 이러한 알고리즘을 구현할 수 있는 의사결정나무분석 솔루션 중의 하나인 Answer Tree 1.0을 사용하였고, 변수의 성격이 범주형 데이터이고 예측변수와 목표변수간의 관계를 찾아야 할 때 가장 유용한 방법인 CHAID(Chi-squared Automatic Interaction Detection)방법을 적용하였다. 환자의 특성 및 진료내역 중 어떠한 특성이 재원기간을 결정짓는데 가능성이 높은가를 예측하려 할 경우, CHAID알고리즘은 최상의 예측변수로서 결정된 변수를 이용하여 재원기간을 결정짓는데 가장 큰 차이를 갖는 두 개 이상의 구분된 집단으로 나누고 그 결과를 트리구조로 나타낸다¹³⁾. CHAID알고리즘의 카이제곱 통계량을 분리기준으로 사용하여 p값이 지정한 유의수준보다 크면, 그 예측변수(환자의 특성 및 진료내역)는 목표변수(재원기간)의 분류에 영향을 주지 않는 것으로 간주하여 차식 마디를 형성할 대상에서 제외하였다. 이와 같이 재원기간 적정군과 비적정군의 특성을 도출하였으며, 이를 규칙화하여 제시하였다.

4. 데이터 마이닝 결과 분석 및 모형의 예측력 평가

의사결정나무를 형성한 후에는 형성된 나무가 얼마나 타당성을 가지고 만들어졌는지를 평가하는 것이 중요한데, 이는 각 마디에 대한 타당성을 평가하는 이익 도표(gain chart)와 오분류 확률(misclassification probability)을 나타내는 위험 도표(risk chart)를 가지고 평가하였다. 이에 따라 재원기간 적정군 및 비적정군을 결정짓는 규칙을 실제 환자 데이터에 적용시킴으로써 모형의 예측력을 평가하였다. 이는 실제 환자의 데이터 값과 데이터 마이닝 모형에 의해 내려진 결정이 일치하는 가를 평가하기 위한 방법이다. 환자의 분류 즉, 재원기간 적정군과 비적정군에 대한 예측분류 값을 가지고 평가할 수 있다. 즉 실제 적정군을 모형에서 적정군이라고 분류한 값인 민감도(sensitivity)와 실제 비적정군을 모형에서 비적정군이라고 분류한 특이도(specificity) 및 예측도(predictive rate)를 산출함으로써 데이터마이닝 모형의 예측력을 평가하였다¹⁴⁾. 또한 데이터 마이닝 결과 도출된 적정군을 결정짓는 규칙과 기 개발된 CP와의 비교 검증을 실시하여 그 일치 여부에 consistent rule과 inconsistent rule로 구분하였고 inconsistent rule 중 acceptable rule을 설정¹⁵⁾하여 데이터 마이닝 기법으로부터 도출된 규칙에 대한 타당성 검증을 실시하였다.

III. 연구 결과

1. 연구대상자의 특성

연구 대상인 제왕절개술 환자 261명을 대상으로 주요 항목별 분포를 살펴 본 결과는 표 1과 같다. 환자의 연령층은 30-34세가 113명으로 43.3%, 25-29세가 74명으로 28.4%의 순을 나타내었으며, 40세 이상의 노령임산부는 18명으로 6.9%의 비율을 나타냈다. 임신주수는 38-40주가 69.3%로 가장 많았고, 분만 적절 시기인 36주 이상의 비율이 234명으로 89.7%, 비적절 시기인 36주 미만의 비율이 27명으로 10.3%를 나타냈으며, 임신주수와 상관성이 높은 신생아 체중은 215명인 82.4%가 정상 체중의 범주에 속했다. 그리고 제왕절개술에 대한 일반적인 적응증

4개 항목(반복 제왕절개술, 난산으로 인한 분만진행의 실패, 둔위, 태아곤란증)에 대한 분포는 반복 제왕절개술이 116명인 44.4%로 가장 높은 비율을 나타냈고, 난산으로 인한 분만진행의 실패가 9명으로 3.4%, 태아곤란증이 5명으로 1.9%를 나타냈다. 또한 항생제 투여는 penicillins계 투여가 134명으로 51.3%, penicillins계와 cephalosporines계의 복합 투여가 106명으로 40.6%, cephalosporines계 투여가 21명으로 8.1%의 순을 나타냈다.

2. 데이터 마이닝 기법으로부터 도출된 재원기간 결정 규칙

그림 2와 같이 재원기간의 분류와 예측의 과정을 나무구조에 의한 추론규칙에 의해 표현하였다. 전체 환자 261명 중 재원기간이 6인 이하인 적정군은 198명으로 전체 261명의 75.9%, 재원기간이 6일 초과인 비적정군은 63명으로 24.1%를 나타냈다. penicillin계를 단독 투여했을 때 적정군의 비율이 85.8%로 penicillin

Table 1. Descriptive statistics for the study sample

Category	Measure	Value	Count	Percent
Basic Information	Age group	20-24	7	2.7
		25-29	74	28.4
		30-34	113	43.3
		35-39	49	18.8
		40-44	16	6.1
		45-	2	0.8
	Intrauterine period (weeks)	<33	17	6.5
		34-35	10	3.8
		36-37	43	16.5
		38-40	181	69.3
41-42		10	3.8	
Birth weight (kg)	Under weight (<2.5)	33	12.6	
	Normal (2.5 ≤ 4.0)	215	82.4	
	Over weight (>4.0)	13	5.0	
Laboratory test	Hematocrit (%)	Normal (≥30)	157	60.2
		Abnormal (<30)	104	39.8
Indication	Repeat procedure	Yes	116	44.4
	Dystocia	Yes	9	3.4
	Breech presentation	Yes	-	-
	Fetal distress	Yes	5	1.9
Medication	Antibiotics	Penicillins	134	51.3
		Cephalosporins	21	8.1
		Penicillins&Cephalosporins	106	40.6
Complication	Puerperal Infection	Yes	-	-
	Puerperal pyrexia	Yes	-	-
	Venous complication	Yes	-	-
	Obstetric embolism	Yes	-	-
	Complication of anaesthesia	Yes	-	-
	Disruption of caesarean - section wound	Yes	2	100.0
Outcome	Lengths of stay (days)	Appropriate group (≤6)	198	75.9
		Inappropriate group(> 6)	63	24.1

계와 cephalosporin계를 복합 투여했을 때의 75.5%와 cephalosporin계를 단독 투여했을 때의 14.3% 보다 각각 1.1배, 6배 높게 나타났다. 환자의 특성 및 진료내역 중 재원 기간을 결정짓는 가장 중요한 특성은 항생제 종류로, 재원기간을 결정짓는 두 번째 특성은 penicillin계를 단독 투여군과 cephalosporin계 단독 투여군에 따라 달라짐을 알 수 있었다. penicillin계를 단독 투여한 군에서 재원기간을 결정짓는 가장 중요한 특성은 수술시의 임신주수로 임신주수가 35주이하인 군에서는 적정군이 62.5%, 35주 초과인 군에서는 89.0%로 임신주수가 35주 초과인 군에서의 적정군의 비율이 그렇지 않은 군에 비해 1.4배 높게 나타났다. cephalosporin계를 단독 투여한 군에서 재원기간을 결정짓는 가장 중요한 특성은 hematocrit수치로

hematocrit 수치가 29이하인 군에서는 적정군이 42.9%, 29초과인 군에서는 0%로 hematocrit수치가 29 이하인 군에서의 적정군의 비율이 그렇지 않은 군에 비해 매우 높게 나타남을 알 수 있었다. 그리고 이외에도 신생아체중, 연령 등이 재원기간에 영향을 미치는 중요한 변수임을 알 수 있었다.

그리고 표 2와 같이 재원기간 적정군과 비적정군 별 환자의 특성 및 진료내역을 규명하고 이를 결정짓는 규칙을 생성하였다. 또한 이와 같은 예측 분류에 대한 신뢰도를 측정함으로써 예측가능성을 점수화하였다. 적정군을 결정짓는 '규칙 1' 이 경우 '임신주수가 24주에서 35주, 신생아체중인 2.2kg에서 3.0kg, 투여 항생제의 종류가 penicillin계' 인 경우 재원일수 비열외군일 확률이 100%임을 나타낸다. 비적정군을

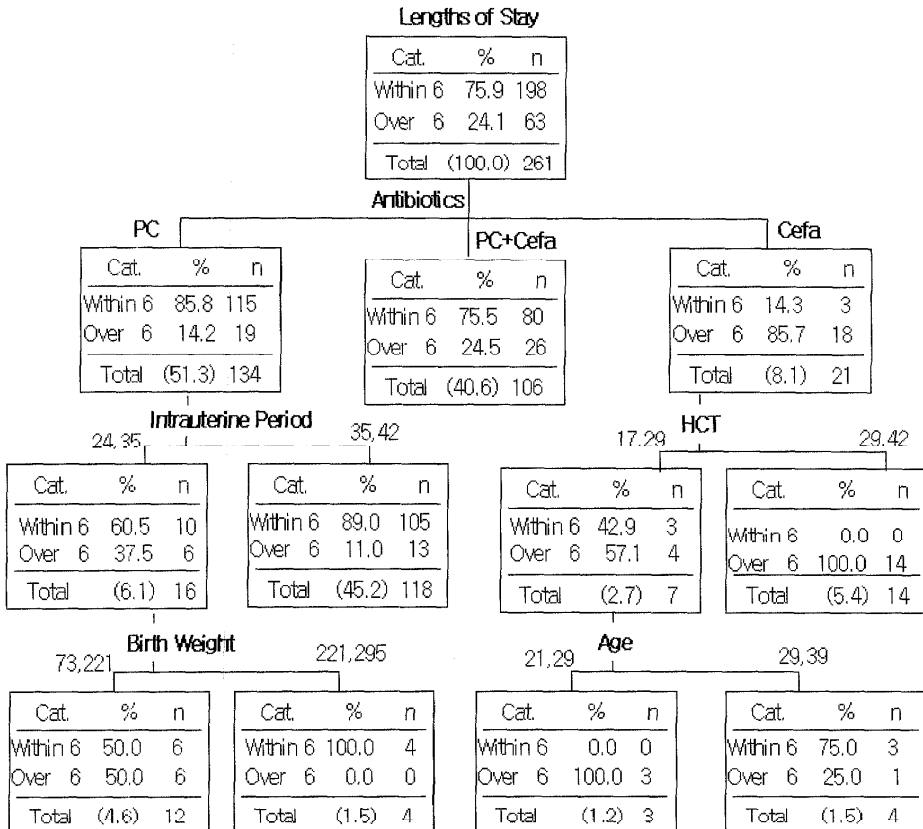


Fig 2. Tree structure for the prediction of lengths of stay

결정짓는 '규칙 1'의 경우 'hematocrit 수치가 29에서 42, 투여 항생제의 종류가 cephalosporin계'인 경우 재원일수 열외군일 확률이 100%임을 나타낸다.

3. 데이터 마이닝 결과 분석 및 모형의 예측력 평가

실제 환자의 데이터 값과 데이터 마이닝으로부터 추출된 저장규칙에 의한 값이 일치하는지를 평가한 결과, 표 3과 같이 true positive와 true negative를 합한 값인 예측도는 82.4%, true positive인 민감도는 100.0%, true negative인 특이도는 27.0%를 나타냈다.

그리고 데이터 마이닝 결과 도출된 재원기간 적정군을 결정짓는 규칙과 기개발된 CP와의 비교 검증을 실시하여 consistent rule과 inconsistent rule로 구분하고 inconsistent rule 중 acceptable rule을 설정하여 데이터 마이닝 기법으로부터 도출된 규칙에 대한 타당성 검증을 실시한 결과는 다음과 같다.

기존 CP와 데이터 마이닝에 의해 새롭게 도출된 규칙이 정확하게 일치되는 consistent rule은 없었고 대부분이 inconsistent rule이었지만 비교 가능한 항목들이 있었다. 기존 CP에서는 수술 후의 항생제 투여에 대하여 투여 여부만을 제시하고 있지만 데이터

마이닝에서 도출된 적정군 결정 규칙에서는 penicillin계 투여가 재원기간을 적절하게 할 수 있는 유의한 변수로 도출됨을 알 수 있었다.

이에 따라 향후 CP개발시 기존 개발된 CP를 보완하여 환자의 여러 특성에 따른 적절한 치료 내역을 명시화 할 수 있음을 알 수 있다. 또한 기존 CP에서는 제왕절개수술 환자에 대해 hematocrit 검사의 실시만을 제시하고 있고 이의 결과에 따른 진료 내역은 제시하고 있지 않지만, 데이터 마이닝 결과에서는 환자의 연령 및 hematocrit결과 수치와 관련되어 어떠한 항생제를 쓸 경우 적정 재원일수를 갖게 될지를 예측할 수 있었다. 또한 임신주수, 신생아체중 등 환자의 여러 특성과 관련되어 재원일수를 예측할 수 있는 규칙을 도출할 수 있었다.

4. 데이터 마이닝 기법을 활용한 CP 개발 방안

지속적인 임상과정 향상을 위한 CP의 개발과 관리 위해서 CP의 개발 및 실행 후 이의 적용 결과를 측정하고 분석하여 개선을 지속시키기 위한 활동이 이루어져야 하며, 이 때 데이터 마이닝 기법을 적용시킴으로써 과거의 축적된 데이터 기반의 객관적이고 명시적인 CP개발을 도모할 수 있다. 데이터

Table 2. Decision rules for the prediction of lengths of stay

Lengths of stay	Rule	Age	Intrauterine period	Birth weight	HCT	Antibiotics	Reliability of prediction
Appropriate group (within 6 days)	1	.	24 ≤ Intrauterine period ≤ 35	2.2 (Birth weight ≤ 3.0)	.	PC *	100.0%
	2	.	35 < Intrauterine period ≤ 42			PC	89.0%
	3	PC+Cefa	75.5%
	4	29 < Age ≤ 39	.	.	17 ≤ HCT ≤ 29	Cefa	75.0%
Inappropriate group (over 6 days)	1	.	.	.	29 < HCT ≤ 42	Cefa **	100.0%
	2	21 ≤ Age ≤ 29	.	.	17 ≤ HCT ≤ 29	Cefa	100.0%

* PC: Penicillin, ** Cefa: Cephalosporin

Table 3. Prediction rate for the CHAID model for lengths of stay

		Actual		Total	Correct Prediction Rate
		Inappropriate	Appropriate		
Predicted	Inappropriate	17 (27.0%) **	0 (0%)	17 (6.5%)	82.4%
	Appropriate	46 (73.0%)	198 (100.0%) *	244 (93.5%)	
	Total	63 (100.0%)	198 (100.0%)	261 (100.0%)	

*: sensitivity, **: specificity

마이닝 기법은 다음과 같이 CP개발 및 관리 상에 활용될 수 있다.

첫째, CP에 포함되어야 할 진료과정을 결정하는 과정에 데이터 마이닝 기법을 활용할 수 있다. CP의 개발을 위해서는 대상 환자군과 CP의 적용 범위를 결정한 뒤, 목표를 설정하여 기존 문헌과 연구결과를 검토하여 목표 달성 방법에 대한 가설을 설정하여 진료 자료를 수집하고 환자의 진료결과자료를 분석하여야 한다. 이로써 환자평가, 검사, 투약, 처치, 교육 상의 진료 내용의 포함 여부 및 제공 시점 및 횟수 등을 결정해야 한다. 이 때 진료 결과를 목표변수로 하고 이들의 내용을 예측변수로 포함하는 데이터 마이닝 모형을 설정할 수 있다. 예를 들어 제왕절개 수술 환자의 진료 과정에서의 항생제 투여의 경우, 만약 항생제 A의 투여 여부를 결정하려고 한다면, 데이터 마이닝 분석 결과를 통해 항생제 A를 투여한 군과 그렇지 않은 군에 따른 진료결과를 분석하여 적정군과 비적정군의 특성을 도출하여 규칙화함으로써, 이의 CP에의 포함 여부를 결정할 수 있을 것이다. 마찬가지로 여러 종류의 항생제 중 최상의 결과를 이끌 수 있는 항생제의 종류를 선택하고자 할 경우도, 각 항생제의 종류에 따른 진료 내역 및 진료결과를 규명하고 이를 규칙화하며, 이와 같은 예측 분류에 대한 신뢰도를 측정하여 예측가능성을 점수화 할 수 있다. 이와 같은 방법으로 투약 및 처치의 제공 시점 및 횟수에 따른 진료 내역 및 결과를 추론규칙에 의해 표현하고 계량화할 수 있다. 이와 같이 도출된 지식들은 궁극적으로 시간-업무 교차표(time-task matrix)이 형태를 지니는 CP이 규격화된 진료 내역 작성시 진료 내용의 구성요소들과 시간적 순서의 결정에 활용될 수 있다.

둘째, 중증도, 동반질환 등 환자의 특이성을 반영하는 진료 내역 결정에 활용될 수 있다.

CP가 각 질환 또는 처치에 대한 가장 보편성을 지닌 표준이라고 할 수 있지만, 같은 진단군의 환자라도 보편적인 특성 외의 특이성을 가지고 있어 표준적인 CP의 적용이 어려운 경우, 융통성 있는 접근이 이루어져야 하며, 이 때 이에 대한 적절한 진료 내역을 제공할 수 있는 객관적이고 타당한 근거가 있어야 한다. 데이터 마이닝 기법을 통해 진료 내역 및 결과에 영향을 미치는 환자의 특이성을 도출할 수

있으며, 특이성에 따른 적정 진료 내역을 분석을 통해 도출할 수 있다. 이로써 진료 의사는 일률적인 CP의 활용이 아닌, 축적된 환자 데이터에 근거한 환자의 특성을 최대한 반영한 임상적 의사 결정을 내릴 수 있다.

셋째, 변이(variance)분석에 활용될 수 있다. 변이자료의 분석결과는 진료결과를 반영하는 것이다. 합병증 발생률, 치료 지연율, 감염률 등의 지표는 진료결과를 반영하는 변이로서 진료과정 관리의 필요성을 반영하는 것이다¹⁵⁾. 데이터 마이닝 기법은 재원기간 동안 진료과정이 CP에서 벗어나거나, 기대되었던 결과가 이루어지지 않은 상태 즉 합병증 발생, 치료지연 등의 진료 결과가 발생되었을 경우, 이의 원인 파악 및 개선안 설정에 활용될 수 있다. 진료결과의 질을 결정짓는 규칙을 도출함으로써 기존 설정된 CP에 대하여 적절성 여부 검토하며 향후 CP의 재조정 등에 활용할 수 있다.

IV. 고 찰

일부 질환 또는 수술에 대한 의료보험 진료비의 포괄수가제 적용은 병원 경영이라는 측면뿐만 아니라 환자 진료 양상에 있어서 매우 큰 변화를 요구하고 있다. 과거에 정부는 진료 수가 인상을 억제하고 의료보험 진료비 심사 기능을 통하여 국민의료비의 증가를 통제하여 왔지만, 앞으로는 포괄수가제를 통하여 진료비를 지불하는 동시에 진료의 질적 수준을 모니터링하는 정책을 도입할 것으로 예상된다¹⁶⁾. 이러한 의료 환경에 능동적으로 대처하기 위해서 병원은 진료 비용 절감과 의료의 질적 수준 유지라는 두 가지 목표를 달성하기 위한 노력이 필요하다. CP는 의료의 질적 수준 유지 및 향상과 진료 비용 절감을 얻을 수 있는 유용한 수단이 될 수 있을 것이다.

CP는 진료과정과 진료자원들을 효율적으로 배치함으로써 진료의 질을 높이고 불필요한 검사나 치료의 중복을 피하게 함으로써 재원일수를 줄여 비용을 절감할 수 있는 방안중의 하나이다. 그러나 일단 개발된 CP라 할지라도 지속적으로 개선할 수 있는 부분을 찾아내 고치고 적용해가는 과정을 반복함으로써 최상의 진료서비스를 제공할 수 있어야 한다¹⁷⁾. 즉 한국의 현실과 의료 환경, 가장 최근의 자료를 토대

로 근거 중심의 의료서비스 제공하는 것이 필요하다. 자원일수 단축에 있어서 환자의 임상적 특성과 함께 의료진의 의료서비스 제공 패턴의 변화가 가장 중요한 요인으로 제시됨에 따라, 의료기관에서는 비용효과적이고 과학적인 CP를 개발하여 병원경영의 합리화를 추구하며 본 연구에서 분석한 것처럼 의료의 질을 유지하면서 불필요한 의료자원 소모량을 감소시켜 나가는 전략적 설정이 필요하다. 향후 DRG 제도 하에서는 의료기관간의 경쟁이 심화되고 환자들은 보다 높은 양질의 서비스를 요구함에 따라 의료기관은 정확한 데이터를 수집하여, 데이터에 기초한 분석 기법을 이용하여 의사결정 정보를 활용하는 정보체계의 구축이 필요하다⁷⁾. 미국의 센트라헬스센터에서 데이터 마이닝 기법을 이용하여 폐렴 환자에 대한 CP를 개발한 결과 환자당 2,000 달러의 비용 절감효과를 본 결과 등을 볼 때 데이터 마이닝 기법과 같은 정보기술을 이용하여 CP를 개발하는 것은 매우 의미가 있는 것을 알 수 있다⁷⁾.

본 연구에서는 의료분야의 의사결정지원을 위한 지식경영체제 도입을 촉진시키기 위한 방안으로써 데이터 마이닝 기법을 적용하여 대규모 데이터 내에 존재하지만 숨겨져 있는 상호관련성과 패턴에 대한 탐색을 통해 유용한 지식을 이끌어 내고자 하였다. 의료분야에서의 데이터 마이닝 응용 분야로는 의료 이용도 분석, 삭감률 분석, 질병 패턴 분석, 진료과정 및 결과 분석, 건강증진 관련 분석, 경영 분석 등이 있을 수 있다. 본 연구에서는 진료과정 및 결과 분석의 일환으로 제왕절개수술환자의 특성과 진료 내역, 이에 따른 재원기간을 예측할 수 있는 지식을 발굴하였으며, 데이터 마이닝 기법을 활용한 제왕절개수술 환자를 위한 CP의 개발 방안을 제시하였다.

그러나 본 연구에도 몇 가지 제한점이 있었다. 충분한 양의 데이터를 추출하지 못했으며, 데이터 필드의 부재 등으로 인하여 진료내역 분석에 어려움이 있었다. 재원 기간의 예측을 위해 설정된 예측변수 즉, 제왕절개수술 환자의 특성 및 치료에서 중요한 특성으로 파악되어지는 항목들의 선정은 임상진료지침, 임상의학의 견해, 실제 환자 데이터에 대한 기술통계량 측정을 통한 데이터의 특성 파악을 통해 이루어져야 한다. 이로써 환자의 특성에 따른 처치 및 투약 내역, 합병증, 재원기간 간의 상관관계 및 패턴을

동시에 파악하여 이를 CP개발에 활용할 수 있다. 그러나 본 연구가 전향적 연구가 아닌 과거 데이터를 기반으로 한 후향적 연구로 이루어짐에 따라 데이터 필드의 부재와 중복 등으로 인하여 다자원 분석에 필요한 데이터 획득의 어려움이 있었다. 환자의 특성에 대한 세부 속성들이 더 포함되었어야 하며, 또한 시술 및 치료 과정에서 각종 검사, 주사, 처치, 투약 등에 세부 정보가 포함되었어야 한다. 이러한 변수들이 충분히 모형에 포함되지 못한 제한점이 있었다. 또한 선정된 변수를 포함한 모형의 분석 결과에서도, 주요한 요인으로 예측했던 제왕절개 분만 환자의 적응증이 분석 모형에서 유의한 요인으로 도출되지 않았고, 임신주수, 연령, 신생아체중, hematocrit 만이 환자의 특성 중 재원기간에 영향을 미치는 유의한 요인으로 도출되었다. 환자의 특성 및 진료내역에 관한 데이터를 더 수집할 수 있었다면 데이터 마트의 구조에 상세한 속성들을 추가로 포함시킴으로써 상대적으로 중요한 속성을 가려내고, 데이터마트의 양적, 질적으로 우수한 데이터에 비해하는 데이터 마이닝의 적중률을 향상시킬 수 있었을 것이다. 그리고 보다 많은 사례수를 포함하여 분석하였다면 보다 나은 모형의 설명력과 신뢰도를 얻을 수 있었을 것이다. 또한 데이터 마이닝 기법으로부터 도출된 규칙에 대한 타당성 검증을 위해 데이터 마이닝의 분석 대상이 아닌 다른 환자를 대상으로 발굴된 새로운 규칙을 적용시키고 이의 결과를 측정하여 보다 타당한 시스템의 유용성 평가 과정을 거치지 못한 제한점이 있었다.

하지만 본 연구는 과거의 축적된 환자 데이터를 이용하여 데이터 마이닝이란 새로운 기법을 사용함으로써 기존의 조사나 연구에서 발견할 수 없었던 의미 있는 새로운 상관관계, 패턴, 추세를 밝혀내고 이와 같은 유용한 정보를 집적하고 체계화하여 환자의 치료 및 관리를 위한 CP에 활용할 수 있는 근거를 제시한다는데 그 의의가 있다고 할 수 있다.

향후에는 데이터 마이닝 기법을 통해 발굴된 지식을 활용하여 CP를 개발하고 실제 진료에 적용시켜 가치를 재평가하는 중재적 연구가 필요하며, 운영 데이터의 축적·활용을 통한 지식의 검증 및 보완이 이루어져야 할 것이다. 또한 적용 결과를 분석하여 이의 유용성을 평가함으로써 임상에서의 지식 기반의

임상진료지침 개발의 활성화를 도모해야 할 것이다.

본 연구 결과에서 데이터 마이닝 기법을 통해 도출된 재원기간 결정 규칙과 모형의 예측력이 아직 임상에 적용되기에는 미흡한 부분이 있지만, 임상이론과 전문의, 그리고 이에 의해 축적된 정보인 환자 데이터를 이용한 데이터 마이닝 기법을 효율적으로 활용한 CP를 개발하여 그 타당성을 점차적으로 높여감으로써 실용화 단계에 이를 수 있도록 하여야 할 것이다.

참고문헌

1. Jung HW, Mun HS, Park EH, Sim EY, Oh JY, Lee SJ et al. Development and clinical application of critical pathway for the Cesarean Section. Proceedings of the conference of the Korean Society of Medical Quality Assurance 1999:339-354
2. Steven D, Sharon F, Jane R, Cook E, Francis SD, Thomas H. Critical Pathways Intervention to Reduce Length of Hospital Stay. The American Journal of Medicine 2001;110(3):175-180
3. Health Insurance Review Agency. The Annual Report of Statistics, Seoul, Korea, 1999
4. Chung CS. Clinical application of critical pathway with an example of stroke. Korean journal of Medical Quality Assurance 1999;6(1):12-17
5. Giuse DA, Miller NB. Strategies for medical knowledge acquisition. Medical informatics, Houten, Springer, 1997
6. Sung won Cho. Data Mining Approach and Case Studies Based on Transformation Consulting Methodology. Proceedings of the conference of the Korean Society of Intelligent Information System 1998:130-137
7. Daelim Information & Telecommunication, Knowledge Management Committee. This is Core of Knowledge Management, Seoul, Korea, Changhae, 1998
8. Cho NC, Sohn MH, Kim TH, Lee HS. Implementing A Data Warehouse Metadata : A Case. Proceedings of the conference of the Korean Society of Intelligent Information System 1999:383-392
9. Fayyad U, Piatetskyo G, Smyth P. Advances in Knowledge Discovery and Data Mining, Menlo Park, Calif, AAAI Press, 1996
10. Ho SH, Chae YM, Cho KW, Jee SH, Lee DH. Data Mining Application for Knowledge Management in Medical Field, Korean Journal of medical informatics 1999;5(3):169-179
11. The Korean Society of Obstetrics and Gynecology. Obstetrics, 3th ed, Seoul, Korea, 1997
12. Choi JH, Han ST, Kang HC, Kim ES. Data mining Decision Tree Analysis using Answer Tree, Seoul, Korea, SPSS Academy, 1998
13. Piatetsky G, Fayyad WJ. Advances in knowledge discovery and data mining, AAAI/MIT Press, 1996
14. Ho SH, Chae YM, Cho SY, Choi DH, Song YW, Park CS. et al. Development of decision support system for the management of hypertension using data mining technology. Proceedings of the conference of the Korean Society of Intelligent Information System 2000:271-282
15. Korea Health Industry Development Institute. Critical Pathway, Seoul, Korea, 2001
16. Lee SI. Overview of critical pathway for its successful development and implementation in Korea. Korean journal of Medical Quality Assurance 1999;6(1):6-11
17. Choi MA. A study of KDRG classification system and factors related to length of stay using data mining methodology : for the cesarean section. Graduate school of health science and management, Seoul, Korea, Yonsei University, 2001