

## 머신러닝을 이용한 급성 뇌졸중 퇴원 환자의 중증도 보정 사망 예측 모형 개발에 관한 연구

백설경<sup>1</sup>, 박종호<sup>2</sup>, 강성홍<sup>3</sup>, 박혜진<sup>4\*</sup>

<sup>1</sup>아주대학교병원, <sup>2</sup>계명대학교 동산의료원, <sup>3</sup>인제대학교 보건행정학과, <sup>4</sup>대구가톨릭대학교 국제의료경영학과

## A study on the development of severity-adjusted mortality prediction model for discharged patient with acute stroke using machine learning

Seol-Kyung Baek<sup>1</sup>, Jong-Ho Park<sup>2</sup>, Sung-Hong Kang<sup>3</sup>, Hye-Jin Park<sup>4\*</sup>

<sup>1</sup>Ajou University Medical Center

<sup>2</sup>Keimyung University Dongsan Medical Center

<sup>3</sup>Dept. of Health Policy & Management, Inje University

<sup>4</sup>Dept. of International Healthcare Administration, Daegu Catholic University

**요약** 본 연구는 머신러닝을 활용하여 급성 뇌졸중 퇴원 환자의 중증도 보정 사망 예측 모형 개발을 목적으로 시행하였다. 전국 단위의 퇴원손상심층조사 2006~2015년 자료 중 한국표준질병사인분류(Korean standard classification of disease-KCD 7)에 따라 뇌졸중 코드 I60-I63에 해당하는 대상자를 추출하여 분석하였다. 동반질환 중증도 보정 도구로는 Charlson comorbidity index(CCI), Elixhauser comorbidity index(ECI), Clinical classification software(CCS)의 3가지 도구를 사용하였고 중증도 보정 모형 예측 개발은 로지스틱회귀분석, 의사결정나무, 신경망, 서포트 벡터 머신 기법을 활용하여 비교해 보았다. 뇌졸중 환자의 동반질환으로는 ECI에서는 합병증을 동반하지 않은 고혈압(hypertension, uncomplicated)이 43.8%로, CCS에서는 본태성고혈압(essential hypertension)이 43.9%로 다른 질환에 비해 가장 월등하게 높은 것으로 나타났다. 동반질환 중증도 보정 도구를 비교해 본 결과 CCI, ECI, CCS 중 CCS가 가장 높은 AUC값으로 분석되어 가장 우수한 중증도 보정 도구인 것으로 확인되었다. 또한 CCS, 주진단, 성, 연령, 입원경로, 수술유무 변수를 포함한 중증도 보정 모형 개발 AUC값은 로지스틱 회귀분석의 경우 0.808, 의사결정나무 0.785, 신경망 0.809, 서포트 벡터 머신 0.830로 분석되어 가장 우수한 예측력을 보인 것은 서포트 벡터머신 기법인 것으로 최종 확인되었고 이러한 결과는 추후 보건의료정책 수립에 활용될 수 있을 것이다.

**Abstract** The purpose of this study was to develop a severity-adjustment model for predicting mortality in acute stroke patients using machine learning. Using the Korean National Hospital Discharge In-depth Injury Survey from 2006 to 2015, the study population with disease code I60-I63 (KCD 7) were extracted for further analysis. Three tools were used for the severity-adjustment of comorbidity: the Charlson Comorbidity Index (CCI), the Elixhauser comorbidity index (ECI), and the Clinical Classification Software (CCS). The severity-adjustment models for mortality prediction in patients with acute stroke were developed using logistic regression, decision tree, neural network, and support vector machine methods. The most common comorbid disease in stroke patients were hypertension, uncomplicated (43.8%) in the ECI, and essential hypertension (43.9%) in the CCS. Among the CCI, ECI, and CCS, CCS had the highest AUC value. CCS was confirmed as the best severity correction tool. In addition, the AUC values for variables of CCS including main diagnosis, gender, age, hospitalization route, and existence of surgery were 0.808 for the logistic regression analysis, 0.785 for the decision tree, 0.809 for the neural network and 0.830 for the support vector machine. Therefore, the best predictive power was achieved by the support vector machine technique. The results of this study can be used in the establishment of health policy in the future.

**Keywords** : Stroke, Medical service quality outcome, Mortality rate, Machine learning, Prediction model, Severity-Adjustment, Comorbidity

본 논문은 한국연구재단 연구과제로 수행되었음(NRF-2018R1D1A1B07050009).

\*Corresponding Author : Hyejin Park(Daegu Catholic Univ.)

Tel: +82-53-850-2593 email: hjpark@cu.ac.kr

Received August 13, 2018

Revised October 17, 2018

Accepted November 2, 2018

Published November 30, 2018

## 1. 서론

뇌졸중(stroke)은 갑작스러운 뇌혈류 장애로 발생하는 신경학적 증상 및 징후가 24시간 이상동안 지속되거나 사망에 이르게 하는 경우를 총칭하며 크게 뇌경색(cerebral infarction)과 뇌출혈(cerebral hemorrhage)로 인한 뇌졸중으로 분류할 수 있다. 뇌졸중은 단일 질환으로는 우리나라 사망률 1위의 질환일 뿐만 아니라 사망하지 않은 경우라도 인지기능 손상, 상하지 마비, 안면마비, 언어장애, 연하곤란 및 식이장애, 시각장애, 만성통증, 감각이상 등 심각한 장애를 초래하여 삶의 질을 저하시키고 가족에게는 심리적 부담과 함께 사회경제적인 부담까지 증가시키고 있다[1-3].

우리나라의 경우 급속한 고령화로 인하여 빠른 속도로 만성질환이 증가하고 있으며 특히 뇌졸중과 같은 중증도 질환은 재난적 의료비를 증가시켜 국민을 중산층에서 빈곤층으로 전락시킬 수 있는 가능성이 높아 국가적으로 예방 정책 수립과 관리가 필요하다고 할 수 있을 것이다.

최근 수준 높은 의료서비스를 제공하고자 하는 국민의 요구와 함께 의료기술의 발달, 의료에 대한 소비자의 접근성 확대, 의료기관의 급속한 증가와 대형화는 의료의 질 경쟁을 더욱 심화시키고 의료서비스를 제공받은 환자의 건강결과(outcome)에 대한 평가를 점차 중시하게 되는 계기가 되었다. 이러한 변화는 양질의 의료서비스 제공에 대한 필요성을 점차 증대시켰고 대부분의 의료서비스를 제공하는 의료기관은 양질의 의료서비스 제공이 궁극적 목표가 되었다.

양질의 의료서비스에 대한 평가는 일반적으로 구조(structure), 과정(process), 결과(outcome) 영역에서 시행된다. 그 중 과정평가는 서비스 제공자들이 구체적인 질 향상을 위한 행동 지침 활용이 쉽다는 이점이 있으나 결과와의 연관성 증명과 함께 세부적인 행동단위로 기술되어 있기 때문에 많은 비용이 든다는 단점이 있다[4].

반면에 결과평가는 평가내용과 기준에 대한 합의가 쉽고 간결하여 의료서비스 질 평가에서 중요시되고 있다[5]. 제공된 의료서비스의 질에 대한 결과 평가는 사망률, 재입원율, 유병률, 합병증 발생률, 재수술률, 장애 발생률 등을 통해 주로 검증된다. 가장 일반적으로 사용되는 결과지표인 퇴원 환자 사망률은 제공된 의료서비스의 질을 최종적으로 측정하며, 의료기관에서 제공하는 의료

서비스의 안정성, 적시성, 효율성, 형평성 등을 평가할 수 있는 중요한 지표로 활용되고 있다[6].

그러나 환자 사망에 대한 결과 평가에 대한 검증은 우선 중증도 보정이 선행되어야 한다. 즉, 퇴원환자 사망에 대한 비교 결과는 환자의 중증도가 서로 다르기 때문에 환자의 중증도가 반영되지 못하면 결과에 대한 타당성이 확보되지 못하여 의료기관 간 또는 국가 간의 비교가 어려울 수 있다. 이러한 의미로 볼 때 중증도를 보정한 퇴원 시 사망률은 의료서비스 질 평가 비교 측정을 위해 사용되는 대표적 지표라고 할 수 있다[6].

주진단에 동반된 동반질환(comorbidity)은 주진단의 사망에 영향을 미치게 된다. 동반질환이란 질환의 주요 원인과 직접적으로 관련되지 않은 질환으로 사망 혹은 의료자원 사용의 증가를 가져올 수 있는 증상 또는 진료 결과에 영향을 주는 주진단명과는 다른 질병을 의미한다[7]. 동반질환은 결과에 직접 영향을 줄 수 있을 뿐 아니라 환자 건강상태를 결정하는 주요한 요인으로 환자의 부담을 증가시킬 수 있다. 동반질환을 고려하지 않고 결과를 측정하거나 예후를 예측하는 것은 결국 왜곡된 결과를 초래할 수 있다. 따라서 주진단의 중증도와 함께 환자의 동반질환에 대한 위험을 체계적으로 보정하여야만 진료의 질적 차이와 올바른 결과 평가 및 예측이 가능하므로 동반질환의 보정은 매우 중요하다고 볼 수 있다.

동반질환 보정방법으로는 Elixhauser comorbidity index(ECI), Kaplan-Feinstein 점수, Charlson comorbidity index(CCI)등이 있다[7-9]. 특히 이러한 보정방법 중 다수의 선행연구에서 동반질환을 이용하여 중증도를 보정하는 방법으로 ECI와 CCI를 많이 활용하였다. ECI와 CCI는 이미 신뢰도와 타당도가 충분히 검증된 바 있어 중증도 보정 사망모형 개발 등 의료의 질 평가에서 널리 활용되고 있다.

최근 2016년 다보스 세계경제포럼에서 4차 산업혁명을 처음 언급한 이후 4차 산업혁명에 대한 전 세계적으로 관심이 점차 증대되고 있다. 4차 산업혁명의 핵심 내용 중 빅데이터(big data)를 기반으로 한 인공지능(artificial intelligence)에 대한 관심이 매우 높고 또한 이러한 관심은 보건의료분야 연구에까지 확대되어 최근 인공지능을 활용하고 그 결과를 적용하려는 노력들이 점차 증가하고 있다. 앞으로는 다양한 형태의 빅데이터를 분석하여 미래를 예측하고 보다 새롭고 진보된 가치를 창출함으로써 그 결과를 기반으로 보건의료정책 수립에 보

템이 될 수 있을 것으로 사료된다.

이제까지 뇌졸중 환자 중증도 보정 사망과 관련된 대부분의 선행 연구는 동반질환에 대한 각각의 중증도 보정 도구를 활용하여 살펴보았기 때문에 다양한 보정 방법 중 어떠한 방법이 더 나은 예측력을 제공하는지 알기 어려웠다.

이에 본 연구에서는 동반질환 중증도 보정 방법들을 서로 비교하여 살펴본 후 어떤 도구를 사용하였을 때 더 나은 예측력을 제시 할 수 있는지 확인해 보고자 한다. 또한 최근 각광을 받고 있는 머신러닝을 이용하여 급성 뇌졸중 퇴원 환자의 중증도를 보정한 사망 예측 모형을 개발하고 적절한 보정방법에 대한 가이드라인을 제시하고자 한다.

## 2. 연구방법

### 2.1 연구대상

#### 2.1.1 급성 뇌졸중 입원환자 정의

건강보험심사평가원의 급성 뇌졸중 환자 정의에 따라 수집된 자료 중 주진단 코드가 출혈성 뇌졸중(I60-I62)과 허혈성 뇌졸중(I63)에 해당되는 I60, I61, I62, I63을 급성 뇌졸중 환자로 정의하였다.

#### 2.1.2 분석 대상

본 연구를 위해 질병관리본부로부터 2006년부터 2015년까지의 퇴원손상심층조사 자료 2,084,004건을 수집하였다.

퇴원손상심층조사는 전체 입원환자를 대상으로 조사되었으며 조사항목으로는 환자 인적사항에 해당되는 성, 연령, 거주지, 보험유형 등 항목과환자의 주진단, 부진단, 수술, 입·퇴원일, 입원경로, 재원일수 등 환자의 진료와 관련된 정보가 포함되는 전국 단위의 자료이다[10]. 또한 질병관리본부에서 미국의 퇴원환자 조사(National Hospital Discharge Survey, NHDS)를 기반으로 하여 구축된 것으로 급성기 일반병원 중 시도별, 병상규모별 표본병원을 선정하여 매년 조사된 자료이다[11-12].

2006년부터 2015년까지 10년간의 퇴원손상심층조사 자료 2,084,004건 중 주진단이 한국표준질병사인분류(Korean standard classification of disease-KCD 7) 코드 I60-I63인 급성 뇌졸중 퇴원환자 37,266건을 추출하였

고, 이 중 입원경로가 기타, 불명에 해당하는 50건을 제외한 37,216건을 급성 뇌졸중 환자의 중증도 보정 사망 예측 모형개발을 위한 분석대상자로 추출하였다. 우선 2006년부터 2015년까지 10년간의 퇴원손상심층조사 자료 중 2005, 2007, 2009, 2011, 2013년 자료는 모형개발 및 내적 타당도 검증 그룹으로 하여 분석대상자는 19,088건으로 선정하였고, 모형의 과적합(overfitting)문제 해결을 위한 외적타당도 검증용으로 2006, 2008, 2010, 2012, 2014년 자료를 활용하여 18,128건으로 최종 선정하였다.

### 2.2 변수정의

Table 1. Definition of variables

Variables	Definition
Sex	Male, Female
Age	<49, 50-64, 65-74, 75+
Admission route	Emergency room, Ambulatory
Principal diagnosis code	The diagnosis established after study to be chiefly responsible for occasioning the patient's episode of care in hospital (or attendance at the health care facility)[13]
Additional diagnosis code	A condition or complaint either coexisting with the principal diagnosis or arising during the episode of care or attendance at a health care facility[13]
History of surgery	Yes, No

### 2.3 동반질환 측정도구

본 연구에서는 동반질환 측정을 위한 도구로 Elixhauser comorbidity index(ECI), Charlson comorbidity index(CCI), clinical classification software(CCS)를 사용하였다. CCI는 17개 질병에 대하여 1~6점까지의 일정한 가중치를 부여한 뒤 이 가중치의 합을 보정하는 방법으로 '0, 1, 2, 3+'의 점수로 범주화시켜 평가하는 도구이다[14-15].

ECI는 31개 동반상병에 대해 만성질환 입원환자 동반상병 유무를 산출하며 이분형태 그대로 분석에 사용한 다[16].

CCS는 미국 AHRQ(The Agency for Healthcare Research and Quality)에 의해 개발된 질병을 임상학적으로 군집화한 방법으로써 질병군 분류방법으로 널리 활용되는 도구이다[17-18].

### 2.4 분석방법

모형의 타당성 및 정확성 확보를 위해 분석 대상자들

모형개발 및 내적 타당도 검증 그룹과 외적 타당도 검증을 위한 두 그룹으로 분류하여 분석하였다. 우선 퇴원손상심층조사 자료 중 2005, 2007, 2009, 2011, 2013년 자료는 모형개발 및 내적 타당도 검증 그룹으로 하였고, 모형의 과적합(overfitting)문제 해결을 위한 외적타당도 검증용으로 2006, 2008, 2010, 2012, 2014년 자료를 활용하였다.

빈도 분석을 실시하여 대상자의 일반적 특성, 주진단 분포, CCI, ECI, CCS에 따른 동반질환 분포를 파악하였다.

동반질환 중증도 보정 사망 예측 모형 개발을 위한 분석기법은 로지스틱 회귀분석(logistic regression model), 의사결정나무(decision tree model), 신경망(neural network model), 서포트 벡터 머신(support vector machine model)을 이용하였고, 모형 개발 후 모형의 적합성 평가 및 선정을 위한 예측력은 AUC(area under the curve of ROC)값으로 평가하였다. AUC는 0.5에서 1사이 값을 갖는데 0.5정도인 경우 예측력이 없고 0.8 이상일 때 좋은 예측력이 있음을 의미하며, 1인 경우 완전한 예측력을 있다고 할 수 있다. 모든 자료 분석은 KNIME software를 이용하였다.

### 3. 연구결과

#### 3.1 연구대상자의 특성

##### 3.1.1 연구대상자의 일반적 특성

본 연구의 총 대상인원은 37,216명이었고 대상자의 일반적 특성을 살펴보면 [Table 2]와 같다.

Table 2. General characteristics of acute stroke inpatients

Characteristics		Model development & internal validation		External validation	
		N	%	N	%
Gender	Male	10,284	53.9	9,822	54.2
	Female	8,804	46.1	8,306	45.8
Age	-49	2,598	13.6	2,494	13.8
	50-64	5,543	29.0	5,142	28.4
	65-74	5,602	29.3	5,273	29.1
	75+	5,345	28.0	5,219	28.8
Admission route	Emergency room	14,421	75.6	13,909	76.7
	Ambulatory	4,667	24.4	4,219	23.3
History of surgery	No	16,282	85.3	15,391	84.9
	Yes	2,806	14.7	2,737	15.1
	All	19,088	100.0	18,128	100.0

모형개발 및 내적 타당도 검증 그룹에서 성별은 남성이 53.9%로 여성 46.1%에 비해 조금 더 많았으며, 연령 분포는 49세 이하를 제외한 50-64세, 65-74세, 75세 이상에서 비슷한 분포를 보였다. 입원경로는 응급실을 통해서 입원한 경우가 75.6%로 외래 24.4%에 비해 높았다. 수술 유무는 수술하지 않은 경우가 85.3%로 수술한 경우 14.7%에 비해 높게 나타났다(Table 2).

##### 3.1.2 대상자의 주진단 분포

모형개발 및 내적 타당도 검증 그룹에서 지주막하 출혈(subarachnoid hemorrhage, I60)이 10.1%, 뇌내출혈(intracerebral hemorrhage, I61)이 16.5%, 비외상성 두개내 출혈(other nontraumatic intracranial hemorrhage, I62) 4.0%, 뇌경색(cerebral infarction, I63) 69.4%로 나타나 뇌경색이 가장 높았다[Table 3].

Table 3. Distribution of principal diagnosis

Diagnosis	Model development & internal validation		External validation	
	N	%	N	%
Subarachnoid hemorrhage(I60)	1,919	10.1	1,875	10.3
Intracerebral hemorrhage(I61)	3,151	16.5	2,860	15.8
Other nontraumatic intracranial hemorrhage(I62)	770	4.0	749	4.1
Cerebral infarction(I63)	13,248	69.4	12,644	69.7
All	19,088	100.0	18,128	100.0

##### 3.1.3 대상자의 동반상병 분포

모형개발 및 내적 타당도 검증 그룹에서 동반상병 중증도 보정의 대표적 방법 중 하나인 CCI에 따른 분포 결과는 (Table 4)와 같다. 0점에 해당되는 경우가 50.2%로 가장 많았고 1점이 28.6%, 2점 12.4%, 3+ 8.8% 순으로 나타났다(Table 4).

Table 4. Distribution of Charlson comorbidity index

Score	Model development & internal validation		External validation	
	N	%	N	%
0	9,584	50.2	9,154	50.5
1	5,456	28.6	5,217	28.8
2	2,365	12.4	2,185	12.1
3+	1,683	8.8	1,572	8.7
All	19,088	100.0	18,128	100.0

또한 모형개발 및 내적 타당도 검증 그룹 분석 대상자 중 ECI에 따른 동반상병 분포는 아래 표에서와 같이 합병증 동반하지 않은 고혈압(hypertension, uncomplicated)이 43.8%로 월등하게 가장 높았고, 합병증 동반하지 않은 당뇨병(diabetes, uncomplicated) 19.8%, 마비(paralysis)가 9.5%, 심부정맥(cardiac arrhythmias)이 9.0% 순으로 높았다(Table 5).

**Table 5.** Distribution of comorbidity disease by Elixhauser comorbidity index

Disease	Model development & internal validation		External validation	
	Yes		Yes	
	N	%	N	%
AIDS/HIV	2	0.0	1	0.0
Alcohol abuse	99	0.5	98	0.5
Blood loss anemia	4	0.0	6	0.0
Cardiac arrhythmias	1,727	9.0	1,706	9.4
Chronic pulmonary disease	290	1.5	325	1.8
Coagulopathy	93	0.5	93	0.5
Congestive heart failure	389	2.0	392	2.2
Deficiency anemia	84	0.4	90	0.5
Depression	266	1.4	280	1.5
Diabetes, complicated	495	2.6	491	2.7
Diabetes, uncomplicated	3,784	19.8	3,545	19.6
Drug abuse	0	0.0	0	0.0
Fluid and electrolyte disorders	169	0.9	157	0.9
Hypertension, complicated	35	0.2	36	0.2
Hypertension, uncomplicated	8,360	43.8	7,786	43.0
Hypothyroidism	113	0.6	120	0.7
Liver disease	269	1.4	267	1.5
Lymphoma	12	0.1	13	0.1
Metastatic cancer	87	0.5	74	0.4
Obesity	4	0.0	2	0.0
Other neurological disorders	602	3.2	619	3.4
Paralysis	1,809	9.5	1,747	9.6
Peptic ulcer disease excluding bleeding	116	0.6	145	0.8
Peripheral vascular disorders	186	1.0	166	0.9
Psychoses	21	0.1	27	0.1
Pulmonary circulation disorders	65	0.3	67	0.4
Renal failure	392	2.1	397	2.2
Rheumatoid arthritis/collagen vascular diseases	56	0.3	65	0.4
Solid tumor without metastasis	314	1.6	290	1.6
Valvular disease	247	1.3	185	1.0
Weight loss	14	0.1	14	0.1

모형개발 및 내적 타당도 검증 그룹에서 CCS에 따른 동반상병 분포는 (Table 6)과 같다.

본태성 고혈압(essential hypertension)이 43.9%, 합병증 없는 당뇨병(diabetes mellitus without complication)이 19.7%, 급성 뇌혈관 질환(acute cerebrovascular

disease)이 11.4%, 마비(paralysis) 9.9% 순으로 높은 것으로 나타났다.

**Table 6.** Distribution of comorbidity disease by clinical classification software category

Disease	Model development & internal validation		External validation	
	Yes		Yes	
	N	%	N	%
Diabetes mellitus without complication	3,764	19.7	3,640	20.1
Diabetes mellitus with complications	536	2.8	476	2.6
Disorders of lipid metabolism	1,803	9.4	1,705	9.4
Deficiency and other anemia	394	2.1	338	1.9
Senility and organic mental disorders	785	4.1	707	3.9
Affective disorders	307	1.6	240	1.3
Other hereditary and degenerative nervous system conditions	504	2.6	467	2.6
Paralysis	1,892	9.9	1,707	9.4
Other nervous system disorders	1,062	5.6	1,025	5.7
Essential hypertension	8,385	43.9	7,758	42.8
Coronary atherosclerosis and other heart disease	614	3.2	562	3.1
Cardiac dysrhythmias	1,731	9.1	1,638	9.0
Acute cerebrovascular disease	2,179	11.4	2,205	12.2
Other and ill-defined cerebrovascular disease	766	4.0	663	3.7
Late effects of cerebrovascular disease	869	4.6	771	4.3
Pneumonia (except that caused by tuberculosis or sexually transmitted disease)	703	3.7	655	3.6
Aspiration pneumonitis; food/vomitus	324	1.7	323	1.8
Gastritis and duodenitis	650	3.4	475	2.6
Other gastrointestinal disorders	405	2.1	413	2.3
Chronic renal failure	306	1.6	338	1.9
Urinary tract infections	514	2.7	447	2.5
Hyperplasia of prostate	361	1.9	328	1.8
Residual codes; unclassified	1,183	6.2	1,184	6.5

주) 전체 259개 질환 중 모형개발 및 내적검증 질환의 수가 300건 이상인 질환을 주요 동반질환으로 하여 일반적 특성을 제시함. 모형개발 시에는 전체 259개 질환 모두를 포함함.

### 3.2 중증도 보정 사망 예측 모형 개발

#### 3.2.1 로지스틱 회귀분석을 이용한 중증도 보정 사망 예측 모형 개발

##### 3.2.1.1 모형 개발

Fig. 1에서와 같이 KNIME software로 로지스틱 회귀 분석으로 모형을 구축하였다.

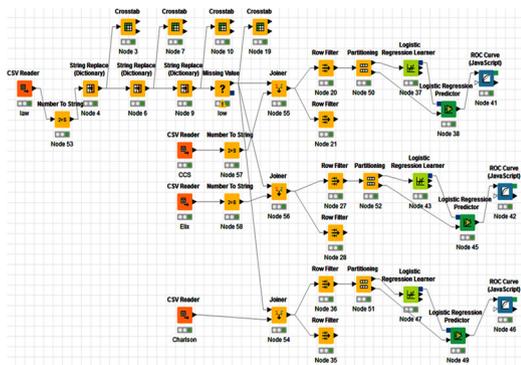


Fig. 1. Logistic regression model development using KNIME

3.2.1.2 모형 평가

로지스틱 회귀분석을 이용하여 중증도 보정 모형을 개발하고 평가하였다. 우선 CCI, ECI, CCS 각각에 대한 모형은 CCS에서 AUC값 0.750로 가장 높았다.

두 번째 단계로 주진단까지 포함된 결과에서도 CCS가 가장 높은 AUC 값을 나타냈고 세 번째 단계인 성별과 연령까지 포함하였을 때, 입원경로와 수술유무 변수까지 포함한 네 번째 단계에서도 각각 0.798과 0.808로 높게 나타났다.

로지스틱 회귀분석을 이용하였을 때에는 CCS, 주진단, 성, 연령, 입원경로, 수술유무 변수를 포함한 경우가 AUC 값 0.808로 가장 높은 모형 예측력을 보였다. 또한 대부분 AUC 값 대부분 0.5보다 높아 예측 모형의 성능과 안정성이 유효함을 확인하였다(Table 7).

Table 7. Logistic regression model assessment using AUC

Model		Internal validation AUC	External validation AUC
step1	CCI	0.546	0.529
	ECI	0.634	0.647
	CCS	0.750	0.760
step2	CCI + PD	0.700	0.705
	ECI + PD	0.744	0.759
	CCS + PD	0.788	0.799
step3	CCI + PD + Gender + Age	0.734	0.741
	ECI + PD + Gender + Age	0.764	0.777
	CCS + PD + Gender + Age	0.798	0.807
step4	CCI + PD + Gender + Age + AR + ES	0.758	0.766
	ECI + PD + Gender + Age + AR + ES	0.785	0.796
	CCS + PD + Gender + Age + AR + ES	0.808	0.817

CCI : Charlson comorbidity index  
 ECI : Elixhauser comorbidity index  
 CCS : clinical classification software  
 PD : principal diagnosis  
 AR : admission route  
 ES : history of Surgery

3.2.1.3 모형의 결과

또한 로지스틱 회귀분석을 통해 뇌졸중 환자의 중증도 보정 사망률 예측 모형을 개발한 결과 심정지 및 심실세동(cardiac arrest and ventricular fibrillation) 상관 계수가 3.746으로 가장 높았고 오심 및 구토(nausea and vomiting) 3.144, 외인에 의한 폐질환(lung disease due to external agents) 2.799 순으로 높은 것으로 나타났다 [Fig. 2][Table 8].

Table 8. Severity-adjusted mortality rate model for acute stroke patients using logistic regression

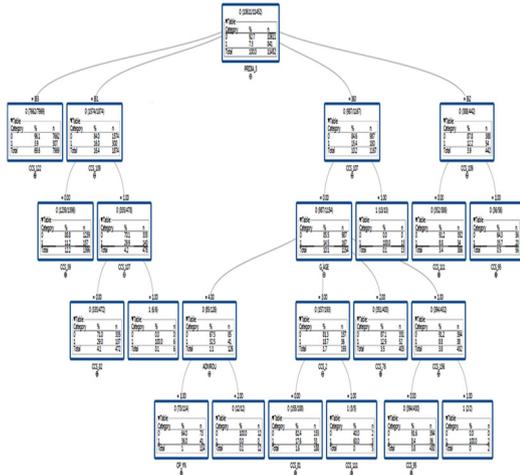
Variable	Coeff.	Std. Err.	z-score	P> z	
Constant	-2.323	0.160	14.527	0.000	
Principle Diagnosis	I61	0.217	0.129	-1.679	0.093
	I62	-0.418	0.206	2.030	0.042
	I63	-1.106	0.146	7.586	0.000
Age	50-64	-0.141	0.136	1.036	0.300
	65-74	0.062	0.140	-0.442	0.658
	75+	0.636	0.140	-4.536	0.000
Admission route	Ambulatory	-0.974	0.144	6.775	0.000
History of surgery	Yes	0.248	0.114	-2.184	0.029
Septicemia(except in labor)	Yes	2.537	0.363	-6.983	0.000
Cancer of head and neck	Yes	2.194	1.038	-2.114	0.035
Cancer of stomach	Yes	1.552	0.561	-2.766	0.006
Cancer of breast	Yes	3.189	1.217	-2.621	0.009
Cancer of prostate	Yes	1.767	0.835	-2.116	0.034
Leukemias	Yes	2.370	0.685	-3.458	0.001
Other endocrine disorders	Yes	1.514	0.544	-2.784	0.005
Coagulation and hemorrhagic disorders	Yes	0.866	0.404	-2.145	0.032
Encephalitis(except that caused by tuberculosis or sexually transmitted disease)	Yes	2.139	0.688	-3.109	0.002
Coma; stupor; and brain damage	Yes	2.397	0.700	-3.424	0.001
Other nervous system disorders	Yes	0.673	0.206	-3.267	0.001
Acute myocardial infarction	Yes	2.110	0.343	-6.146	0.000
Cardiac dysrhythmias	Yes	0.370	0.150	-2.472	0.013
Cardiac arrest and ventricular fibrillation	Yes	3.746	0.439	-8.535	0.000
Congestive heart failure; nonhypertensive	Yes	0.739	0.261	-2.830	0.005
Acute cerebrovascular disease	Yes	1.059	0.100	-10.636	0.000
Pneumonia(except that caused by tuberculosis or sexually transmitted disease)	Yes	0.901	0.160	-5.616	0.000
Respiratory failure; insufficiency; arrest (adult)	Yes	2.266	0.732	-3.098	0.002
Lung disease due to external agents	Yes	2.799	1.166	-2.400	0.016
Other lower respiratory disease	Yes	0.879	0.333	-2.641	0.008
Biliary tract disease	Yes	1.181	0.496	-2.382	0.017
Gastrointestinal hemorrhage	Yes	1.026	0.420	-2.444	0.015
Acute and unspecified renal failure	Yes	0.982	0.278	-3.535	0.000
Chronic renal failure	Yes	1.170	0.246	-4.766	0.000
Chronic ulcer of skin	Yes	1.139	0.317	-3.595	0.000
Skull and face fractures	Yes	1.641	0.624	-2.630	0.009
Nausea and vomiting	Yes	3.144	1.314	-2.392	0.017
Residual codes; unclassified	Yes	0.344	0.162	-2.119	0.034

주) 통계적으로 유의한 변수(P>|z|<0.05)만 표로 제시함.

### 3.2.2 의사결정나무를 이용한 중증도 보정 사망 예측 모형 개발

#### 3.2.2.1 모형 개발

Fig. 2에서와 같이 의사결정나무 기법으로 중증도 보정 사망률 예측모형을 구축하였다.



PRDIA3 : Principle Diagnosis (00, 161, 162, 163)  
 AGE : Age (1 = 49, 2 = 50-64, 3 = 65-74, 4 = 75+)  
 ADMROU : Admission Route (1 = Emergency Room, 2 = Ambulatory)  
 CCI : Septicemia/sepsis (0 = No, 1 = YES)  
 CCI107 : Cardiac arrest and ventricular fibrillation (0 = No, 1 = YES)  
 CCI156 : Nephritis, nephrosis, renal sclerosis (0 = No, 1 = YES)

Fig. 2. Severity-adjusted mortality rate model for acute stroke patients using decision tree

#### 3.2.2.2 모형 평가

의사결정나무를 이용한 모형 개발 및 평가 결과 로지스틱 회귀 분석과 동일한 양상을 보였다.

CCI, ECI, CCS 각각에 대한 모형은 CCS에서 AUC 값 0.743로 가장 높았다.

주진단까지 포함된 결과에서도 CCS가 0.789로 가장 높은 값을 나타냈고 세 번째 단계인 성별과 연령까지 포함하였을 때와 입원경로와 수술유무 변수까지 포함한 네 번째 단계에서도 각각 0.744과 0.785로 높게 나타났다.

로지스틱 회귀분석 결과와 마찬가지로 CCS, 주진단, 성, 연령, 입원경로, 수술유무 변수를 포함한 경우가 0.785로 가장 높은 모형 예측력을 보였다. 또한 대부분 AUC 값 대부분 0.5보다 높아 예측 모형의 성능과 안정성이 유효함을 확인하였다(Table 9).

Table 9. Decision Tree model assessment using AUC

Model		Internal validation AUC	External validation AUC
step1	CCI	0.581	0.536
	ECI	0.597	0.591
	CCS	0.743	0.720
step2	CCI + PD	0.727	0.702
	ECI + PD	0.753	0.726
step3	CCS + PD	0.789	0.759
	CCI + PD + Gender + Age	0.723	0.698
	ECI + PD + Gender + Age	0.714	0.735
step4	CCS + PD + Gender + Age	0.744	0.743
	CCI + PD + Gender + Age + AR + ES	0.738	0.735
	ECI + PD + Gender + Age + AR + ES	0.755	0.768
	CCS + PD + Gender + Age + AR + ES	0.785	0.780

CCI : Charlson comorbidity index  
 ECI : Elixhauser comorbidity index  
 CCS : clinical classification software  
 PD : principal diagnosis  
 AR : admission route  
 ES : history of Surgery

### 3.2.3 신경망을 이용한 중증도 보정 사망 예측 모형 개발

#### 3.2.3.1 모형 개발

Fig. 3에서와 같이 KNIME software로 신경망으로 모형을 구축하였다.

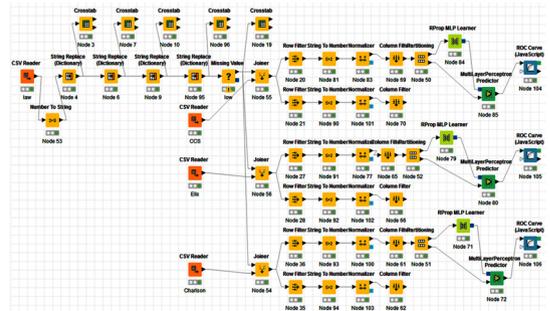


Fig. 3. Neural network model development using KNIME

#### 3.2.3.2 모형 평가

신경망을 이용하여 살펴본 결과 앞서 제시된 로지스틱 회귀분석 및 의사결정나무를 이용한 모형 개발 및 평가 결과와 동일한 양상을 나타내었다.

CCI, ECI, CCS 각각에 대한 모형은 CCS에서 AUC 값 0.752로 가장 높았다.

주진단까지 포함된 결과에서도 CCS가 0.785로 가장 높은 값을 나타냈고 세 번째 단계인 성별과 연령까지 포

함하였을 때와 입원경로와 수술유무 변수까지 포함한 네 번째 단계에서도 각각 0.791과 0.809로 높게 나타났다. 신경망을 이용하여 모형 개발한 결과에서도 CCS, 주진단, 성, 연령, 입원경로, 수술유무 변수를 포함한 경우가 0.809로 가장 높은 모형 예측력을 보였다. 또한 대부분 AUC 값 대부분 0.5보다 높아 예측 모형의 성능과 안정성이 유효함을 확인하였다(Table 10).

Table 10. Neural network model assessment using AUC

	Model	Internal validation AUC	External validation AUC
step1	CCI	0.552	0.529
	ECI	0.635	0.634
	CCS	0.752	0.765
step2	CCI + PD	0.702	0.699
	ECI + PD	0.734	0.745
	CCS + PD	0.785	0.782
step3	CCI + PD + Gender + Age	0.741	0.748
	ECI + PD + Gender + Age	0.763	0.784
	CCS + PD + Gender + Age	0.791	0.794
step4	CCI + PD + Gender + Age + AR + ES	0.761	0.760
	ECI + PD + Gender + Age + AR + ES	0.787	0.798
	CCS + PD + Gender + Age + AR + ES	0.809	0.806

CCI : Charlson comorbidity index  
 ECI : Elixhauser comorbidity index  
 CCS : clinical classification software  
 PD : principal diagnosis  
 AR : admission route  
 ES : history of Surgery

### 3.2.4 서포트 벡터 머신을 이용한 중증도 보정 사망 예측 모형 개발

#### 3.2.4.1 모형 개발

Fig. 4에서와 같이 KNIME software로 서포트 벡터 머신으로 모형을 구축하였다.

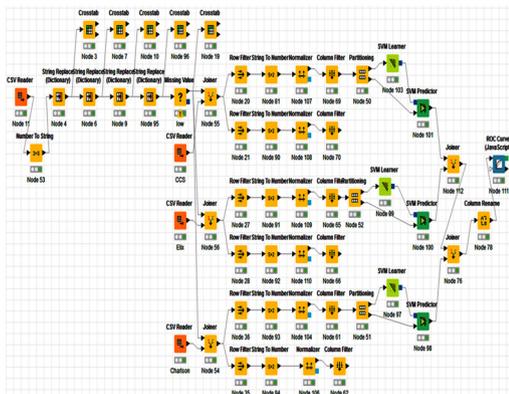


Fig. 4. Support vector machine model development using KNIME

#### 3.2.4.2 모형 평가

마지막으로 서포트 벡터 머신을 이용하여 모형을 개발 및 평가 하였다.

서포트 벡터 머신의 경우도 위의 3가지 방법과 동일한 양상을 보였다.

CCI, ECI, CCS 각각에 대한 모형은 CCS에서 AUC 값 0.748로 가장 높았다.

주진단까지 포함된 결과에서도 CCS가 0.815로 가장 높았으며 성별과 연령까지 포함하였을 때와 입원경로와 수술유무 변수까지 포함한 마지막 단계에서도 각각 0.826과 0.830으로 높게 나타났다.

서포트 벡터 머신을 이용한 경우도 CCS, 주진단, 성, 연령, 입원경로, 수술유무 변수를 포함한 경우가 0.830으로 가장 높은 모형 예측력을 보였다. 또한 대부분 AUC 값 대부분 0.5보다 높아 예측 모형의 성능과 안정성이 유효함을 확인하였다(Table 11).

Table 11. Support vector machine model assessment using AUC

	Model	Internal validation AUC	External validation AUC
step1	CCI	0.570	0.577
	ECI	0.673	0.698
	CCS	0.748	0.766
step2	CCI + PD	0.752	0.752
	ECI + PD	0.742	0.743
	CCS + PD	0.815	0.817
step3	CCI + PD + Gender + Age	0.639	0.560
	ECI + PD + Gender + Age	0.713	0.747
	CCS + PD + Gender + Age	0.826	0.828
step4	CCI + PD + Gender + Age + AR + ES	0.606	0.592
	ECI + PD + Gender + Age + AR + ES	0.807	0.810
	CCS + PD + Gender + Age + AR + ES	0.830	0.837

CCI : Charlson comorbidity index  
 ECI : Elixhauser comorbidity index  
 CCS : clinical classification software  
 PD : principal diagnosis  
 AR : admission route  
 ES : history of Surgery

## 4. 고찰

본 연구는 2006년부터 2015년까지 10년간의 질병관리본부에서 구축한 퇴원손상심층조사 자료를 이용하여 최근 각광 받고 있는 머신러닝 기법을 통해 급성 뇌졸중 퇴원 환자의 중증도 보정 사망 예측 모형을 개발하고자 시행하였다.

급성 뇌졸중 환자는 한국표준질병사인분류(Korean standard classification of disease-KCD 7)에 따라 코드 I60-I63에 해당하는 대상을 추출하여 분석하고 동반질환 중증도 보정 도구로는 Charlson comorbidity index(CCI), Elixhauser comorbidity index(ECI), clinical classification software(CCS)를 이용하였다.

연구 대상자의 성별은 남성이 53.9%로 여성 46.1%에 비해 조금 더 많았으며, 연령분포는 49세 이하를 제외한 50-64세, 65-74세, 75세 이상에서 비슷한 분포를 보였다.

입원경로는 응급실을 통해서 입원한 경우가 75.6%로 외래 24.4%에 비해 높았는데 뇌졸중 질환 특성 상 응급을 요하는 질환 특성으로 인한 결과인 것으로 판단된다. 연구 대상자 대부분 수술하지 않은 경우로 나타났다.

각각의 중증도 보정 도구에 따른 동반질환을 살펴본 결과 CCI에서는 0점이 가장 높았고, ECI에서는 합병증 동반하지 않은 고혈압(hypertension, uncomplicated), 합병증 동반하지 않은 당뇨병(diabetes, uncomplicated), 마비(paralysis), 심부정맥(cardiac arrhythmias) 순으로 높게 나타났다. CCS에 따른 결과로는 본태성고혈압(essential hypertension)이 월등하게 높았고, 합병증 없는 당뇨병(diabetes mellitus without complication), 급성뇌혈관질환(acute cerebrovascular disease), 마비(paralysis) 순으로 높게 나타났다. 본 연구의 ECI와 CCS에 따른 결과에서 동일하게 합병증 동반하지 않은 고혈압과 합병증 동반하지 않은 당뇨병이 가장 높은 것으로 나타났다. 뇌졸중 환자의 동반질환 분포에 대한 이전 논문에서는 출혈성 뇌졸중인 경우 뇌혈관 질환(31.3%), 합병증이 없는 당뇨병(10.1%), 반신마비(9.1%) 등의 순으로 높았고, 허혈성 뇌졸중의 경우는 합병증이 없는 당뇨병(24.2%), 뇌혈관질환(14.3%), 반신마비(8.2%) 순으로 나타났다[19]. 또한 CCI로 살펴본 동반상병 분포에 대한 김남권 등의 연구에서는 허혈성 심질환, 고혈압, 당뇨병 순이었으며, 임지혜 등의 연구에서는 출혈성 뇌졸중 환자 동반질환은 뇌혈관 질환(31.3%), 당뇨병(10.1%), 반신마비(9.1%)였고 허혈성 뇌졸중 환자 동반상병은 합병증이 없는 당뇨병(24.1%), 뇌혈관질환(14.3%), 반신마비(8.2%) 순으로 높은 것으로 나타났다[19]. 이러한 결과는 본 연구의 동반질환 분포 결과와 어느 정도 일치함을 보였다.

3가지 동반질환 중증도 보정 도구 중 CCS인 경우에서 가장 높은 AUC값을 보였으며 성과 연령, 입원경로, 수술유무 변수를 포함하였을 때도 CCS에서 가장 높은

AUC값을 나타내어 가장 우수한 동반질환 중증도 보정 도구인 것으로 알 수 있었다.

선행 연구 결과에 따르면 급성심근경색, 만성폐쇄성 폐질환, 울혈성 심부전, 합병증을 동반한 고혈압, 급성 심혈관질환의 경우 ECI를 이용한 동반질환 보정방법이 CCI 방법보다 사망에 대한 결과 예측력이 월등한 것으로 나타났는데[20-21], 이러한 결과는 본 연구에서도 ECI가 CCI 보다 우수한 예측력을 보인 것과 일치하였다. 동반질환 측정 도구 중 CCI는 1986년에 개발되어 이후 널리 사용된 도구이나 너무 오래 전에 개발되어 의료기술이 빠르게 변화되어가는 현재에 적용함에 한계가 있을 수 있을 것이다. 또한 단순히 4개의 그룹으로 분류하여 가중치로 평가하였기 때문에 세밀하고 정확한 계량화에 대한 신뢰도에도 문제가 있을 수 있다.

아직까지 ECI와 CCS를 예측력을 비교한 연구가 없어 두 가지 도구 중 어떤 것이 더 우수한 도구인지에 대하여 본 연구결과와 비교할 수는 없으나 본 연구결과를 통해 볼 때 3가지 도구 중 CCS가 가장 우수한 예측력을 나타내었고 이러한 결과는 앞으로 사망에 대한 결과 예측 동반질환 보정 도구로 CCS가 가장 적합하다고 할 수 있는 근거로 제시될 수 있을 것으로 사료된다.

본 연구에서는 중증도 보정 사망 예측 모형개발을 위해 머신러닝 기법을 사용하였다. 로지스틱회귀분석, 의사결정나무, 신경망, 서포트 벡터 머신 기법을 이용하여 중증도 보정 사망예측 모형을 개발하였으며 모형의 적합성 측정은 AUC 값으로 평가하였다. 그 결과 가장 우수한 모형 개발 기법은 서포트 벡터 머신인 경우로 나타났다. 이러한 결과는 성, 연령, 입원경로, 수술유무 변수 및 CCS를 포함시켜 분석하였을 시에도 AUC값이 0.830로 가장 높은 값으로 분석되었다.

Hachesu 등은 관상동맥질환, Lily Tapak 등은 당뇨병, Lehmann C등은 알츠하이머 질환 진단 예측에서 여러 알고리즘을 비교한 바 가장 예측력과 정확도가 우수한 것은 서포트 벡터 머신임을 발표하는 등, 여러 선행 연구에서 서포트 벡터 머신이 가장 우수한 예측력을 보임을 제시한 바 있다. 이러한 결과는 본 연구에서도 서포트 벡터 머신이 가장 우수하여 여러 선행 연구들과 동일한 결과를 보임을 확인할 수 있었다[22-24].

본 연구는 전국적 단위의 퇴원환자를 대상으로 하였다는 점과 함께 대부분의 기존 선행 연구에서는 CCI, ECI, CCS 각각의 중증도 보정 지표를 통해 급성 뇌졸중

환자의 동반질환을 살펴보았으나 본 연구에서는 평가 도구를 서로 비교 분석하였다는 점에서 차별성이 있다고 할 수 있겠다.

제한점으로는 본 연구결과가 급성 뇌졸중에 국한되는 사망예측 모형이므로 다른 질환에까지도 동일한 결과를 나타낸다고 제시할 수는 없을 것이다.

따라서 본 연구 결과에서 확장하여 추후에 이러한 결과가 다른 질환의 중증도 보정 사망예측에서도 적용될 수 있을지에 대한 연구가 수행되어야 할 것으로 판단되며, 이러한 연구결과는 앞으로 보건 의료 분야에서 활용될 수 있을 것으로 기대해 본다.

## 5. 결론

퇴원손상심층조사 자료를 통해 전국적 단위의 급성 뇌졸중 퇴원환자를 대상으로 중증도 보정 사망예측 모형을 개발하고 평가하기 위하여 본 연구를 시행하였다.

연구결과 Charlson comorbidity index(CCI), Elixhauser comorbidity index(ECI), clinical classification software(CCS) 3개의 동반질환 보정 도구 중 CCS가 가장 우수한 것으로 나타났으며 로지스틱회귀분석, 의사결정나무, 신경망, 서포트 벡터 머신 기법을 이용한 모형 개발 및 평가한 결과 다양한 기법 중 서포트 벡터 머신 방법이 가장 우수한 설명력을 나타내는 것으로 확인할 수 있었다. 따라서 급성 뇌졸중 환자의 사망 예측을 위한 동반질환 중증도 보정은 CCS를 이용한 서포트 벡터 머신으로 모형을 개발하였을 때 가장 우수한 예측이 가능하다는 것을 확인하였고 이러한 연구결과는 추후 보건 의료 분야 정책 수립에 활용될 것으로 사료된다.

## References

- [1] J. H. Im, K. S. Lee, K. Y. Kim, N. S. Hong, S. W. Lee, H. J. Bae, "Follow-up study on mortality in Korean stroke patients", *J Korean Med Assoc*, Vol. 54 No. 11 pp. 1199-1208, 2011.  
DOI: <https://dx.doi.org/10.5124/jkma.2011.54.11.1199>
- [2] K. J. Kim, M. Heo, I. A. Chun, H. J. Jun, J. S. Lee, H. Jegal, "The re-relationship between stroke and quality of life in Korean adults: based on the 2010 Korean community health survey". *Journal of Physical Therapy Science*, Vol. 27, No. 1, pp. 309-312, 2015.
- [3] E. H. Jeong, O. T. M. S, B. R. Kim, J. Lee, "Relationship between Comorbid Cognitive Impairment and Functional Outcomes in Stroke Patients with Spatial Neglect", *Brain & Neuro Rehabilitation*, Vol. 9, No. 1, 2016.
- [4] H. Rubin, P. Pronovost, G. Diette, "The advantages and disadvantages of process-based measure of health care quality", *International Journal for Quality in Health Care*, Vol. 13, pp. 469-474, 2001.  
DOI: <https://dx.doi.org/10.1093/intqhc/13.6.469>
- [5] J. Y. Yoon, J. Y. Lee, "Development of Outcome Indicators of Urinary Incontinence for Quality Evaluation in Long Term Care Hospitals", *J Korean Acad Nurs*. Vol. 40, No. 1, pp. 110-118, 2010.  
DOI: <https://dx.doi.org/10.4040/jkan.2010.40.1.110>
- [6] J. H. Park, Y. M. Kim, S. S. Kim, W. J. Kim, S. H. K, "Comparison of Hospital Standardized Mortality Ratio Using National Hospital Discharge Injury Data", *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 13, No. 4 pp. 1739-1750, 2012.  
DOI: <https://dx.doi.org/10.5762/KAIS.2012.13.4.1739>
- [7] A. Elixhauser, C. Steiner, D. Harris, R. Coffey, "Comorbidity measures for use with administrative data", *Med Care*, Vol. 36, pp. 8-27, 1998.  
DOI: <https://dx.doi.org/10.1097/00005650-199801000-00004>
- [8] C. Y. Wang, Y. S. Lin, C. Tzao, H. C. Lee, M. H. Huang, W. H. Hsu, H. S. Hsu, "Comparison of Charlson comorbidity index and Kaplan -Feinstein index in patients with stage I lung cancer after surgical resection", *European Journal of Cardio-Thoracic Surgery*, Vol.32, No.1 pp. 877-881, December 2007.  
DOI: <https://dx.doi.org/10.1016/j.ejcts.2007.09.008>
- [9] M. Charlson, P. Pompei, K. Ales, C. MacKenzie, "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation", *J Chronic Dis*, Vol. 40, pp. 373-383, 1987.  
DOI: [https://dx.doi.org/10.1016/0021-9681\(87\)90171-8](https://dx.doi.org/10.1016/0021-9681(87)90171-8)
- [10] S. H. Kang, "Analysis of LOS variance - The results of Korean National Hospital Discharge In-depth Injury Survey 2004~2006", *Public health weekly report, KCDC*, 2, pp. 587-866, 2009.
- [11] Y. M. Kim, "A study on analysis of factors on in-hospital mortality for community-acquired pneumonia", *Journal of the Korean Data & Information Science Society*, Vol. 22, No. 3, pp. 389-400, 2011.
- [12] M. H. Nam, S. S. Kim, I. S. Park, S. H. Kang, W. J. Kim, S. H. Choi, H. K. Jo, Y. T. Kim, S. O. Hong, "A Study on Utilization of non-residential areal in Operation patient", *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 11, No. 6, pp. 2078-2087, 2010.  
DOI: <https://dx.doi.org/10.5762/KAIS.2010.11.6.2078>
- [13] Australia, National Health Data Committee, *National Health Data Dictionary*, Version 12, AIHW, 2003.
- [14] <http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?printer=Y&conceptID=1098>
- [15] J. H. Lee, J. K. Choi, S. N. Jeong, S. H. Choi. "Charlson comorbidity index as a predictor of periodontal disease in elderly participants" *J Periodontal Implant Sci*, Vol. 48, No. 2, pp. 92-102, 2018.
- [16] <http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php>

printer=Y&conceptID=1436

[17] <https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>

[18] Y. H. Choi, Y. J. Kim, S. O. Hong, "Development of Severity-adjusted Length of Stay in Ischemic Heart Disease", Journal of the Korean Data Analysis Society, Vol. 17, No. 1, pp. 407-421, 2015.

[19] K. H. Kim, J. H. Im, "Mortality of Stroke Patients Based on Charlson comorbidity index", Journal of the Korea Contents Association, Vol. 16, No.3, pp. 22-32, 2016. DOI: <https://dx.doi.org/10.5392/JKCA.2016.16.03.022>

[20] J. George, S. Douglas, P. Wagner, F. Alfred, "Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations.", Medical Care, Vol. 39, No. 7, pp. 727-739, 2001.

[21] D. A. Southern, H. Quan, W. A. Ghali, "Comparison of the Elixhauser and Charlson/Deyo Methods of Comorbidity Measurement in Administrative Data", Medical Care, Vol. 42, No. 4, pp. 355-360, 2004. DOI: <https://dx.doi.org/10.1097/01.mlr.0000118861.56848.ee>

[22] P. R. Hachesu, M. Ahmadi, S. Alizadeh, F. Sadoughi, "Use of data mining techniques to determine and predict length of stay of cardiac patients", Health Inform Res Vol. 19, No. 2, pp. 121-9, 2013. DOI: <https://dx.doi.org/10.4258/hir.2013.19.2.121>

[23] L. Tapak, H. Mahjub, O. Hamidi, J. Poorolajal, "Real-Data Comparison of Data Mining Methods in Prediction of Diabetes in Iran", Health Inform Res. Vol. 19, No. 3, pp. 177-185, 2013. DOI: <https://dx.doi.org/10.4258/hir.2013.19.3.177>

[24] C. Lehmann, T. Koenig, V. Jelic, L. Prichep, "Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG)", J Neurosci Methods, Vol. 161, No. 2, pp. 342-350, 2007. DOI: <https://dx.doi.org/10.1016/j.jneumeth.2006.10.023>

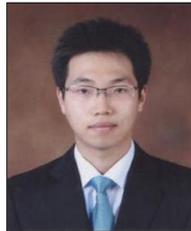
**백 설 경(Seol-Kyung Baek) [정회원]**



- 200년 2월 : 아주대학교 공공정책 대학원 정책학과 보건의료전공
- 2009년 2월 : 아주대학교 의과대학 의료정보학과 박사과정 수료
- 1993년 11월 ~ 현재 : 아주대학교 병원 의무기록팀 팀장

<관심분야>  
보건의료정보

**박 종 호(Jong-Ho Park) [정회원]**



- 2014년 8월 : 계명대학교 경영대학원 의료경영학과 (의료경영학 석사)
- 2018년 8월 : 인제대학교 일반대학원 보건학과 (보건학 박사 수료)
- 1998년 3월 ~ 현재 : 계명대학교 동산의료원 근무

<관심분야>  
보건통계, 데이터마이닝, 인공지능, 정보보호

**강 성 흥(Sung-Hong Kang) [정회원]**



- 1990년 2월 : 서울대학교 보건대학원 보건관리학과 (보건학석사)
- 1997년 2월 : 인제대학교 일반대학원 보건학과 (보건학박사)
- 1998년 3월 ~ 현재 : 인제대학교 보건행정학과 교수

<관심분야>  
보건정보, 의무기록, 데이터마이닝, 건강증진

**박 혜 진(Hye-Jin Park) [정회원]**



- 1999년 8월 : 서울대학교 보건대학원 보건학과 (보건학석사)
- 2012년 8월 : 경북대학교 일반대학원 보건학과 (보건학박사)
- 2013년 9월 ~ 현재 : 대구가톨릭대학교 국제의료경영학과 교수

<관심분야>  
보건의료정보, 공중보건, 건강증진