

머신러닝을 이용한 급성심근경색증 환자의 퇴원 시 사망 중증도 보정 방법 개발에 대한 융복합 연구

백설경¹, 박혜진², 강성홍³, 최준영⁴, 박중호^{5*}

¹아주대학교병원 팀장, ²대구가톨릭대학교 국제의료경영학과 교수, ³인제대학교 보건행정학과 교수,
⁴청암대학교 병원의료정보과 교수, ⁵계명대학교 동산의료원 부주사

Convergence Study in Development of Severity Adjustment Method for Death with Acute Myocardial Infarction Patients using Machine Learning

Seol-Kyung Baek¹, Hye-Jin Park², Sung-Hong Kang³, Joon-Young Choi⁴, Jong-Ho Park^{5*}

¹Ajou University Hospital Team leader

²Dept. of International Healthcare Administration, Daegu Catholic University Professor

³Dept. of Health Policy & Management, Inje University Professor

⁴Dept. of Hospital health information, Cheongam College Professor

⁵Kyeimyong University Dongsan Medical Center Assistant manager

요 약 본 연구는 기존 동반질환을 이용한 중증도 보정 방법의 제한점을 보완하기 위해 급성심근경색증 환자의 맞춤형 중증도 보정방법을 개발하고, 이의 타당성을 평가하기 위해 수행되었다. 이를 위하여 질병관리본부에서 2006년부터 2015년까지 10년간 수집한 퇴원손상심층조사 자료 중 주진단이 급성심근경색증인 한국표준질병사인분류(KCD-7) 코드 I20.0~I20.9의 대상자를 추출하였고, 동반질환 중증도 보정 도구로는 기존 활용되고 있는 CCI(Charlson comorbidity index), ECI(Elixhauser comorbidity index)와 새로이 제안하는 CCS(Clinical Classification Software)를 사용하였다. 이에 대한 중증도 보정 사망예측모형 개발을 위하여 머신러닝 기법인 로지스틱 회귀분석, 의사결정나무, 신경망, 서포트 벡터 머신기법을 활용하여 비교하였고 각각의 AUC(Area Under Curve)를 이용하여 개발된 모형을 평가하였다. 이를 평가한 결과 중증도 보정도구로는 CCS 가 가장 우수한 것으로 나타났으며, 머신러닝 기법 중에서는 서포트 벡터 머신을 이용한 모형의 예측력이 가장 우수한 것으로 확인되었다. 이에 향후 의료서비스 결과평가 등 중증도 보정을 위한 연구에서는 본 연구에서 제시한 맞춤형 중증도 보정방법과 머신러닝 기법을 활용하도록 하는 것을 제안한다.

주제어 : 중증도 보정, 급성심근경색, 동반질환, 머신러닝, 융복합 연구

Abstract This study was conducted to develop a customized severity-adjustment method and to evaluate their validity for acute myocardial infarction(AMI) patients to complement the limitations of the existing severity-adjustment method for comorbidities. For this purpose, the subjects of KCD-7 code I20.0 ~ I20.9, which is the main diagnosis of acute myocardial infarction were extracted using the Korean National Hospital Discharge In-depth Injury survey data from 2006 to 2015. Three tools were used for severity-adjustment method of comorbidities : CCI (charlson comorbidity index), ECI (Elixhauser comorbidity index) and the newly proposed CCS (Clinical Classification Software). The results showed that CCS was the best tool for the severity correction, and that support vector machine model was the most predictable. Therefore, we propose the use of the customized method of severity correction and machine learning techniques from this study for the future research on severity adjustment such as assessment of results of medical service.

Key Words : Severity-adjustment, Acute Myocardial Infarction, Comorbidity, Machine Learning, Convergence study

*Corresponding Author : Jong-Ho Park(jh8283p@naver.com)

Received November 15, 2018

Accepted February 20, 2019

Revised January 24, 2019

Published February 28, 2019

1. 서론

일반적으로 의료의 질은 Donabedian(1983)의 정의에 따라 구조(structure), 과정(process), 결과(outcome) 등 3가지 영역에서 평가될 수 있다[1]. 의료의 질 평가에서 구조(structure)의 평가는 의료의 질을 간접적으로 반영하고, 과정(process)의 평가는 의료 자원 사용의 적정성을 주로 평가하므로 환자의 치료효과를 평가하지 못하는 제한점이 있다. 이에 비하여 결과(outcome)의 평가는 치료효과를 평가하는 것이므로 진정한 의미의 의료의 질 평가라 할 수 있다[2,3]. 결과 평가는 환자마다 질환의 중증도가 다르게 나타나므로 치료결과를 단순 비교하는 것이 아닌 중증도를 보정하여 보는 것이 보다 정확한 결과를 얻을 수 있다. 환자의 중증도 보정은 상당히 복잡한 과정이고 타당성을 보장받기가 용이하지 않음에 따라 결과평가를 위한 중증도 보정방법에 관한 연구에 대한 필요성이 증가하고 있다[2,4]. 이러한 필요성에 부응하여 결과에 영향을 주는 환자의 다양한 특성을 고려하여 위험도를 동등하게 맞추는 중증도 보정 방법에 대한 연구가 이루어져 왔다. 선행연구 결과 환자의 중증도 보정에 영향을 주는 다양한 요인 중 동반상병(comorbidities)이 가장 중요한 예측인자로 판명되었다[2]. 즉, 동반상병이 많은 환자는 치료결과가 좋지 않을 가능성이 더 높기 때문에 결과 평가를 위해서는 동반상병을 이용하여 환자의 중증도를 보정하는 것이 필요하며[4], 이를 위해 Charlson Comorbidity Index(이하 CCI), Elixhauser Comorbidity Index(이하 ECI) 등의 동반상병을 이용한 중증도 보정 방법이 개발되어 결과 평가에 활용되고 있다[2].

그러나 기존에 개발되어 활용되고 있는 동반상병 보정 방법은 두 가지의 제한점이 있다. 첫째는 주진단에 따라 동반상병이 환자의 중증도에 영향을 미치는 정도가 상이함에도 불구하고 기존의 동반상병 보정 방법은 주진단에 상관없이 동일한 방법으로 동반상병을 보정하고 있다. 예를 들어 심장질환 입원환자가 합병증이 없는 당뇨를 동반하고 있는 경우와 단순 목의 손상 입원환자가 합병증이 없는 당뇨를 동반하는 경우에 합병증이 없는 당뇨가 환자의 중증도에 영향을 미치는 정도가 차이가 있음에도 불구하고 기존의 CCI 동반상병 보정 방법은 모두 동일한 가중치인 1점을 부여하여 중증도를 보정하고 있다[5,6]. 둘째는 기존 선행연구에서 건강결과에 영향을 미치는 것으로 나타난 동반상병이 기존 동반상병 보정 방

법에서는 제외된 경우가 있다. 예를 들어 급성심근경색증 입원환자의 경우 심인성 쇼크를 동반한 환자의 사망률이 높은 것으로 나타났음에도 불구하고 기존의 CCI, ECI 동반상병 보정 방법에는 심인성 쇼크 동반이 제외되어 있다[7]. 이와 같이 기존에 개발된 동반상병 보정 방법의 제한점을 보완하는 방법으로, 주진단에 따른 동반상병을 보정하고, 건강결과에 영향을 미치는 동반상병 모두를 고려하는 맞춤형 동반상병 보정 방법을 개발하여 결과 평가의 신뢰도를 높일 필요가 있다.

2016년 다보스포럼에서 제시된 4차 산업혁명(Forth Industrial Revolution)은 이후 전 세계적으로 폭발적인 관심을 끌고 있다[8]. 4차 산업혁명은 인공지능 기술을 통하여 세상을 이해하는 시대이며, 인공지능이란 머신러닝, 딥러닝의 기술을 통해 빅데이터를 분석하고, 예측하여 더 나은 답을 찾아낼 수 있는 기술을 의미한다[9]. 외국의 경우 머신러닝, 딥러닝 기술을 이용하여 의료서비스의 결과 평가에 대한 신뢰성을 향상시키기 위한 연구[10-13]가 활발히 수행되고 있는 반면 우리나라의 경우 이에 관한 연구가 거의 없는 실정이다. 따라서 머신러닝, 딥러닝 등 인공지능 기법을 이용하여 결과 평가의 신뢰성을 향상시킬 수 있는 방안 마련에 대한 국내 연구가 필요하다.

의료의 질을 평가하는 대표적인 결과(outcome) 지표에는 사망률, 진료비, 재원일수 등이 있다[3]. 이 중 사망률 지표는 치료를 통해 개인의 건강상태를 개선한 결과를 나타내는 것으로 안정성, 적시성, 효율성, 형평성 등 의료 질의 다양한 측면을 반영하는 중요한 지표이다[1,4]. 따라서 사망률에 대한 중증도 보정방법의 개발이 우선적으로 필요하다. 통계청의 사망원인통계에 따르면 인구 10만 명당 심장 질환으로 인한 사망률은 2016년 58.2로 각종 암 다음으로 사망률의 순위가 높았다[14]. 대표적인 심장 질환에는 급성심근경색증, 협심증, 심부전, 심내막염 등이 있으며, 이 중 급성심근경색증을 포함한 허혈성 심장질환의 사망률이 가장 높은 비중을 차지하며, 관상동맥이 폐쇄되어 심근 세포가 괴사하는 급성심근경색증은 유병률과 사망률이 지속적으로 증가하고 있음에 따라 사망률 관리가 필요한 질환이다[15]. 모든 질환에 대한 맞춤형 동반상병 보정 방법을 개발하여 결과 평가의 신뢰도를 향상시키는 것이 필요하나 이는 너무 방대한 영역임에 따라 본 연구에서는 급성심근경색증 환자의 퇴원 시 사망에 대한 맞춤형 동반상병 보정 방법 즉, 중증도

보정 방법을 개발하고자 하며, 이를 통해 기존의 중증도 보정방법을 보완할 수 있는 새로운 중증도 보정 방법의 타당성을 제시하고자 한다. 또한 최근 보건 의료에 도입되고 있는 인공지능의 대표적 예측 기법인 머신러닝 기법을 우선적으로 이용하여 맞춤형 동반상병 보정 방법을 개발하고 이를 통해 맞춤형 동반상병 보정 방법의 가능성에 대하여 제시하고자 한다.

2. 연구방법

2.1 연구방법

2.1.1 자료수집

본 연구를 위해 질병관리본부로부터 2006년부터 2015년까지 10년간의 퇴원손상심층조사 자료 2,804,004건을 수집하였다. 퇴원손상심층조사는 의료기관에서 퇴원한 퇴원환자를 조사한 전국단위의 자료이며, 조사항목에는 성, 연령, 입원경로, 주진단, 부진단 등 환자 인적사항 정보와 주요 진료내역 정보가 포함되어 있어 맞춤형 동반상병 보정 방법 즉, 중증도 보정 방법을 개발하기 위한 가장 적합한 자료이다[10].

퇴원손상심층조사자료는 질병관리본부에서 미국의 퇴원환자조사(National Hospital Discharge Survey, NHDS)를 기반으로 하여 구축된 자료로 급성기 일반병원 중 시도별, 병상규모별 표본병원을 선정하여 매년 조사된 자료이다[9,10].

2.1.2 분석대상

건강보험심사평가원의 급성심근경색증 평가대상 질환군의 정의에 따라 주진단이 I21.0~I21.9인 퇴원환자를 급성심근경색증 퇴원환자로 정의하고, 수집된 퇴원손상심층조사 자료 2,804,004건 중 주진단이 I21.0~I21.9인 급성심근경색증 퇴원환자 자료 10,795건을 추출하였다. 이 중 입원경로가 기타 또는 미상인 퇴원환자 자료 9건을 제외한 10,786건의 자료를 분석하였다.

2.2 변수정의

2.2.1 종속변수 및 독립변수 정의

본 연구의 종속변수는 퇴원 시 사망유무이며, 성, 연령, 입원경로, 주진단, 경피적 관상동맥 중재술(PCI) 시행유무, 관상동맥 우회술(CABG) 시행유무, 동반상병 보정

방법 등을 독립변수로 정의하였다. 경피적 관상동맥 중재술(PCI), 관상동맥 우회술(CABG)은 선행연구의 연구방법에 따라 국제 수술 및 시술분류(ICD-9-CM) 코드를 기준으로 입원환자의 수술 및 시술코드 중 36.01~36.09가 부여된 경우 경피적 관상동맥 중재술(PCI) 시행유, 36.10~36.14가 부여된 경우 관상동맥 우회술(CABG) 시행유로 정의하였다[2].

2.2.2 동반상병 보정 방법

본 연구에서는 급성심근경색증 환자의 퇴원 시 사망에 대한 맞춤형 동반상병 보정방법을 개발하였다. 이를 위해 기존의 CCI, ECI 기준을 이용하여 동반상병을 보정하였으며, 선행연구의 연구방법에 따라 Clinical classification software(CCS) 진단군 분류 변수를 이용하여 동반상병을 보정하였다[5-6,16].

CCI는 진료결과에 영향을 미친다고 정의한 17개의 동반상병에 대하여 일정한 가중치를 부여한 뒤 이 가중치의 합(comorbidity score)을 이용하여 환자의 중증도를 보정하는 방법이고, ECI는 진료결과에 영향을 미친다고 정의한 31개의 동반상병에 대해 상병의 동반 여부를 이분형 변수로 처리하여 보정하는 방법이다[5,6]. 본 연구에서 개발한 방법은 AHRQ(The Agency for Healthcare Research and Quality)에 의해 개발된 CCS 기준의 259 질환에 대하여 각 질환별 유무를 독립변수로 하여 보정하는 방법이다. 이는 기존의 ECI 방법처럼 각각의 동반 여부를 이분형 변수로 처리하여 환자의 중증도를 보정하는 방법이다[16]. 기존의 ECI 방법과의 차이는 ECI는 일정 범위의 동반상병만을 독립변수로 하지만 본 연구에서 제안한 방법은 모든 동반상병을 포함하므로 ECI의 제한점인 제외되는 동반상병이 없는 장점을 가지게 된다.

따라서 CCS 진단군 변수를 이용한 동반상병 보정의 경우 입원환자의 모든 동반상병이 포함되어 있음에 따라 급성심근경색증 입원환자의 사망에 영향을 미치는 동반상병이 동반상병 보정 방법에 있어 제외되는 경우는 없다.

2.3 분석방법

2.3.1 머신러닝 분석

정보의 양이 방대해짐에 따라 처리해야 할 데이터의 범위가 광범위해지고 정확하면서 저렴하고 빠른 정보처리와 계산이 이루어지는 머신러닝이 최근 각광받고 있다. 머신러닝은 실시간 분석을 통한 예측 등과 같은 유용한

작업의 수행이 가능하며, 공학 뿐 아니라 의학 등에 활용 및 적용되어 성과를 거두는 등 응용 가능한 범위가 방대해지면서 그 효용성을 인정받고 있다[17]. 머신러닝이란 예측기법을 말한다. 수많은 자료를 읽고 알고리즘을 찾아내어 변화를 예측하는 기술이다[18]. 머신러닝은 학습 데이터를 획득하는 방법에 따라 지도학습(Supervised Learning), 자율학습(Unsupervised Learning) 등으로 구분할 수 있는데, 머신러닝의 지도학습은 입력값과 그에 상응하는 기대 출력값을 데이터로 갖는다. 입력 데이터를 기대 출력값에 최대한 유사하도록 하는 알고리즘이 지도학습 알고리즘이며, 알고리즘은 학습을 실행하였을 때 기대 출력값과 학습된 입력값의 차이를 비교하여 오류를 찾아내며, 이러한 오류는 모델을 수정하는데 근거가 된다[17]. 본 연구에서 개발하고자 하는 급성심근경색증 환자의 퇴원 시 사망에 대한 맞춤형 동반상병 보정 방법 즉, 중증도 보정 방법은 사망 유무라는 학습된 입력값과 사망 예측이라는 기대 출력 값을 기반으로 타당성을 입증하고자 함에 따라 머신러닝의 지도학습을 이용하여 개발하여야 한다. 본 연구에서는 종속변수가 범주형인 경우 적용할 수 있는 머신러닝의 지도학습에 대한 대표적인 알고리즘인 로지스틱 회귀분석(Logistic regression), 의사결정나무(Decision tree), 신경망(Neural network), 서포트 벡터 머신(Support vector machines) 알고리즘[19]을 이용하여 급성심근경색증 환자의 퇴원 시 사망에 대한 중증도 보정 방법을 개발하고자 하며, 이의 타당성을 입증하고자 한다.

로지스틱 회귀분석(Logistic regression)은 통계학에서 오랫동안 사용되어왔던 기법으로, 독립변수들의 조합으로 종속변수의 범주화된 결과값을 모델링하는데 사용된다[20]. 본 연구에서는 반복재가중최소제곱(iteratively reweighted least squares, IRLS) 방법을 이용하여 분석하였으며, 유의한 영향을 미치는 변수만을 결과로 제시하였다. 의사결정나무(Decision tree)는 각 독립변수를 이분화하는 과정을 반복하여 판단 알고리즘을 나무 형태로 형성하는 모델[20]로 주요 분류 기준은 지니 지수(Gini index)를 이용하여 분석하였다. 신경망(Neural network)은 기계에게 인간이나 동물과 같이 학습이 가능케 하고자 뇌의 구조를 수학적으로 모델화한 모형으로 딥 러닝 분석의 기초 알고리즘이다[17]. 신경망 분석은 Profit/Loss 방법을 이용하였다. 서포트 벡터 머신(Support vector machines)은 다차원 공간을 초평면으로 경계를 분류하

는 방법으로 예측력이 높기 때문에 머신러닝 기법 적용 시 많이 적용하는 모형[19]으로 커널함수 중 polynomial Kernel 함수를 이용하여 분석하였다. 분석도구는 오픈소스 머신러닝 분석 도구인 Knime 3.5.3을 이용하여 분석하였다.

2.3.2 모형의 개발

본 연구에서는 머신러닝 분석 기법에 따라 모형을 개발하였으며, 모형개발 시 모델 1은 동반상병 보정, 모델 2는 동반상병, 주진단 보정, 모델 3은 동반상병, 주진단, 성, 연령 보정, 모델 4는 동반상병, 주진단, 성, 연령, 입원 경로, 경피적 관상동맥 중재술(PCI) 시행 유무, 관상동맥 우회술(CABG) 시행 유무를 보정하여 동반상병 보정 방법의 타당성을 평가하였다. 이와 같이 단계별로 변수를 추가하여 모형을 평가한 것은 처음에는 각각의 기법이 동반상병의 보정에 어느 정도 우수함을 보기 위함이고, 단계별로 추가한 것은 실제 중증도 보정에 사용되는 변수를 추가하여 각 기법에서의 중증도 보정의 우수성을 보기 위함이다.

2.3.3 모형의 평가

머신러닝 모델의 목표는 예측이므로, 개발된 모델의 타당성은 전통적으로 민감도와 특이도에 기반한 ROC 곡선(Receiver Operating Characteristic curve)을 이용하여 평가하는 경우가 많다. 이를 위해 데이터를 학습용(training set)과 검증용(test set)으로 분리해야 하며, 과적합을 방지하기 위하여 검증용 데이터는 학습 목적으로 사용해야 한다[20]. 이에 본 연구에서는 선행연구[21] 방법에 따라 2006년부터 2015년까지 수집된 급성심근경색증 퇴원환자 자료 10,795건에 대해 2006년, 2008년, 2010년, 2012년, 2014년 자료는 모형개발 및 내적검증을 위한 자료로 2007년, 2009년, 2011년, 2013년, 2015년 자료는 외적검증을 위한 자료로 활용하였으며, 모형개발 및 내적검증을 위한 자료는 모형개발을 위해 데이터를 학습용 60%, 검증용 40%로 분류하였다.

모형의 ROC 곡선 아래의 면적(area under the curve, 이하 AUC)이 다른 모형의 면적보다 크면 평균적으로 더 우수한 모형이라 할 수 있음[22]에 따라 본 연구에서는 모형의 AUC 값을 이용하여 개발된 모형의 타당성을 검증하였다.

3. 연구결과

3.1 분석대상자의 일반적 특성

3.1.1 분석대상자의 일반적 특성

급성심근경색증 퇴원환자의 일반적 특성을 파악한 결과는 Table 1과 같다. 모델개발 및 내적 검증, 외적검증 모두 성별로는 여성보다 남성의 비율이 높았으며, 연령별로는 50-64세의 비율이 가장 높았다. 입원경로별로는 응급실을 통한 입원이 각각 83.5%, 84.0%로 외래를 통한 입원보다 높았으며, 경피적 관상동맥 중재술(PCI) 시행 유무별로는 시행률이 높은 반면, 관상동맥 우회술(CABG) 시행 유무별로는 시행하지 않음이 월등히 높았다.

Table 1. General characteristics of acute stroke inpatients

Variables		Model Development & Internal Validation		External Validation	
		N	%	N	%
Gender	Male	3,800	69.9	3,728	69.7
	Female	1,635	30.1	1,623	30.3
Age	<49	838	15.4	764	14.3
	50-64	1,803	33.2	1,784	33.3
	65-74	1,456	26.8	1,471	27.5
	75+	1,338	24.6	1,332	24.9
Admission Route	Emergency room	4,540	83.5	4,495	84.0
	Ambulatory	895	16.5	856	16.0
PCI	No	2,247	41.3	2,081	38.9
	Yes	3,188	58.7	3,270	61.1
CABG	No	5,312	97.7	5,251	98.1
	Yes	123	2.3	100	1.9
All		5,435	100.0	5,351	100.0

Footnote) PCI : percutaneous coronary intervention
CABG : coronary artery bypass graft

3.1.2 분석대상자의 주진단 분포

급성심근경색증 퇴원환자의 주진단 분포를 파악한 결과는 Table 2와 같다. 모델개발 및 내적 검증, 외적검증 모두 상세불명의 급성심근경색증(I219)이 각각 40.8%, 40.6%로 가장 높았으며, 급성 심내막하 심근경색증(I214), 전벽의 급성 전층심근경색증(I210) 등의 순으로 높은 것으로 나타났다.

3.2 분석대상자의 동반상병 분포

3.2.1 CCI

급성심근경색증 퇴원환자의 CCI 기준 동반상병 지수

분포를 파악한 결과는 Table 3과 같다. 모델개발 및 내적 검증, 외적검증 모두 0점이 각각 54.8%, 53.8%로 가장 높았으며, 1점, 2점, 3점 이상의 순으로 높은 것으로 나타났다.

Table 2. Distribution of CCI

Variables	Model Development & Internal Validation		External Validation	
	N	%	N	%
I210(Acute transmural myocardial infarction of anterior wall)	696	12.8	678	12.7
I211(Acute transmural myocardial infarction of inferior wall)	615	11.3	539	10.1
I212(Acute transmural myocardial infarction of other sites)	112	2.1	109	2.0
I213(Acute transmural myocardial infarction of unspecified site)	354	6.5	356	6.7
I214(Acute subendocardial myocardial infarction)	1,442	26.5	1,494	27.9
I219(Acute myocardial infarction, unspecified)	2,216	40.8	2,175	40.6
All	5,435	100.0	5,351	100.0

Table 3. Distribution of CCI

Variables	Model Development & Internal Validation		External Validation	
	N	%	N	%
0	2,976	54.8	2,881	53.8
1	1,517	27.9	1,492	27.9
2	586	10.8	597	11.2
3+	356	6.6	381	7.1
All	5,435	100.0	5,351	100.0

3.2.2 ECI

급성심근경색증 퇴원환자의 ECI 기준 동반상병 유무 분포를 파악한 결과는 Table 4와 같다. 모델개발 및 내적 검증, 외적검증 모두 고혈압(Hypertension) 동반이 각각 41.9%, 42.1%로 가장 높았으며, 당뇨(Diabetes, uncomplicated), 심부전(Congestive heart failure) 동반 순으로 높은 것으로 나타났다.

3.2.3 CCS(Clinical classification software)

급성심근경색증 퇴원환자의 CCS 진단군 기준 동반상병 유무 분포를 파악한 결과는 Table 5와 같다. 모델개발

및 내적 검증, 외적검증 모두 본태성 고혈압(Essential hypertension)이 각각 41.9%, 42.1%로 가장 높았으며, 관상동맥죽상경화증 및 기타 심장병(Coronary atherosclerosis and other heart disease) 유, 합병증 없는 당뇨병(Diabetes mellitus without complication) 유 등의 순으로 높은 것으로 나타났다.

Table 4. Distribution of comorbidity disease by ECI

Variables	Model Development & Internal Validation			External Validation		
	Yes		Total	Yes		Total
	N	%	N	N	%	N
AIDS/H1V	2	0.0	5,435	1	0.0	5,351
Alcohol abuse	12	0.2	5,435	16	0.3	5,351
Blood loss anemia	1	0.0	5,435	2	0.0	5,351
Cardiac arrhythmias	352	6.5	5,435	348	6.5	5,351
Chronic pulmonary disease	168	3.1	5,435	167	3.1	5,351
Coagulopathy	17	0.3	5,435	27	0.5	5,351
Congestive heart failure	513	9.4	5,435	527	9.8	5,351
Deficiency anemia	32	0.6	5,435	44	0.8	5,351
Depression	23	0.4	5,435	18	0.3	5,351
Diabetes, complicated	265	4.9	5,435	272	5.1	5,351
Diabetes, uncomplicated	1,306	24.0	5,435	1,308	24.4	5,351
Drug abuse	0	0.0	5,435	0	0.0	5,351
Fluid and electrolyte disorders	45	0.8	5,435	48	0.9	5,351
Hypertension, complicated	25	0.5	5,435	14	0.3	5,351
Hypertension, uncomplicated	2,275	41.9	5,435	2,251	42.1	5,351
Hypothyroidism	32	0.6	5,435	38	0.7	5,351
Liver disease	75	1.4	5,435	89	1.7	5,351
Lymphoma	3	0.1	5,435	4	0.1	5,351
Metastatic cancer	16	0.3	5,435	17	0.3	5,351
Obesity	4	0.1	5,435	3	0.1	5,351
Other neurological disorders	52	1.0	5,435	62	1.2	5,351
Paralysis	15	0.3	5,435	11	0.2	5,351
Peptic ulcer disease excluding bleeding	49	0.9	5,435	44	0.8	5,351
Periphera vascular disorders	88	1.6	5,435	82	1.5	5,351
Psychoses	9	0.2	5,435	3	0.1	5,351
Pulmonary circulation disorders	16	0.3	5,435	12	0.2	5,351
Renal failure	259	4.8	5,435	285	5.3	5,351
Rheumatoid arthritis/collagen vascular diseases	30	0.6	5,435	19	0.4	5,351
Solid tumor without metastasis	87	1.6	5,435	89	1.7	5,351
Valvular disease	105	1.9	5,435	72	1.3	5,351
Weight loss	1	0.0	5,435	1	0.0	5,351

Table 5. Distribution of comorbidity disease by CCS category

Variables	Model Development & Internal Validation			External Validation		
	Yes		Total	Yes		Total
	N	%	N	N	%	N
Diabetes mellitus without complication	1,325	24.4	5,435	1,335	24.9	5,351
Diabetes mellitus with complications	273	5.0	5,435	277	5.2	5,351
Disorders of lipid metabolism	713	13.1	5,435	756	14.1	5,351
Deficiency and other anemia	133	2.4	5,435	154	2.9	5,351
Essential hypertension	2,275	41.9	5,435	2,251	42.1	5,351
Hypertension with complications and secondary hypertension	151	2.8	5,435	145	2.7	5,351
Coronary atherosclerosis and other heart disease	1,498	27.6	5,435	1,525	28.5	5,351
Cardiac dysrhythmias	299	5.5	5,435	286	5.3	5,351
Congestive heart failure; nonhypertensive	402	7.4	5,435	421	7.9	5,351
Late effects of cerebrovascular disease	191	3.5	5,435	175	3.3	5,351
Pneumonia (except that caused by tuberculosis or sexually transmitted disease)	194	3.6	5,435	196	3.7	5,351
Chronic obstructive pulmonary disease and bronchiectasis	115	2.1	5,435	92	1.7	5,351
Other lower respiratory disease	145	2.7	5,435	140	2.6	5,351
Esophageal disorders	106	2.0	5,435	115	2.1	5,351
Gastritis and duodenitis	100	1.8	5,435	106	2.0	5,351
Nephritis; nephrosis; renal sclerosis	102	1.9	5,435	111	2.1	5,351
Acute and unspecified renal failure	152	2.8	5,435	144	2.7	5,351
Chronic renal failure	185	3.4	5,435	221	4.1	5,351
Residual codes; unclassified	700	12.9	5,435	652	12.2	5,351

Footnote) We suggest disease only more than 100 cases that sum model development disease and model verification disease. And we are used total 259 diseases when developing model.

3.3 중증도 보정 사망 예측 모형 개발

3.3.1 로지스틱 회귀분석을 이용한 중증도 보정 사망 예측 모형

3.3.1.1 모형 개발

오픈소스 머신러닝 분석 도구인 KNIME software의 로지스틱 회귀분석으로 중증도 보정 예측 모형을 구축하였다. 데이터 전처리, 모형 개발, 모형 평가 등 로지스틱 회귀분석을 이용한 급성심근경색증 퇴원환자의 동반상병 보정 모형 개발 워크플로어는 Fig. 1과 같다.

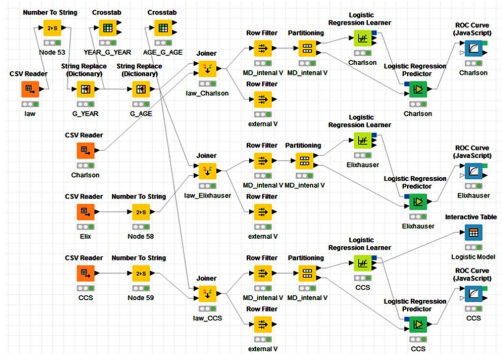


Fig. 1. KNIME workflow for logistic regression model development

3.3.1.2 모형 평가

로지스틱 회귀분석을 이용하여 급성심근경색증 퇴원 환자의 동반상병 보정 모형을 개발하고, AUC 값을 이용하여 개발된 모형을 평가한 결과는 Table 6과 같다. 모델 1, 모델 2, 모델 3, 모델 4 모두 CCS 진단군 분류 변수를 이용한 동반상병 보정 방법이 가장 높은 것으로 나타났다. 모델 1, 모델 2, 모델 3, 모델 4 모형 중 모델 4의 CCS 진단군 분류 변수를 이용한 새로운 동반상병 보정 방법과 주진단, 성, 연령, 입원경로, 경피적 관상동맥 중재술(PCI) 시행 유무, 관상동맥 우회술(CABG) 시행 유무 모형의 내적검증 및 외적검증 AUC값이 각각 0.755, 0.731로 가장 높은 것으로 나타나 이를 이용하여 급성심근경색증 퇴원환자의 중증도 보정 사망 예측 모형을 개발하는 것이 가장 타당한 것으로 평가되었다.

3.3.1.3 모형의 결과

로지스틱 회귀분석을 이용하여 CCS 진단군 분류 변수를 이용한 동반상병 보정 방법과 주진단, 성, 연령, 입원경로, 경피적 관상동맥 중재술(PCI) 시행 유무, 관상동맥 우회술(CABG) 시행 유무를 독립변수로 한 급성심근경색증 퇴원환자의 중증도 보정 사망 예측 모형을 개발

Table 6. Logistic regression model assessment using AUC

Variables		Internal Validation AUC	External Validation AUC
Model1	CCI	0.595	0.585
	ECI	0.610	0.584
	CCS index	0.619	0.594
Model2	CCI + Principle Diagnosis	0.572	0.587
	ECI + Principle Diagnosis	0.595	0.567
	CCS index + Principle Diagnosis	0.652	0.618
Model3	CCI + Principle Diagnosis + Gender + Age	0.616	0.602
	ECI + Principle Diagnosis + Gender + Age	0.708	0.682
	CCS index + Principle Diagnosis + Gender + Age	0.711	0.698
Model4	CCI + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.583	0.570
	ECI + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.699	0.707
	CCS index + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.755	0.731

Footnote) PCI : percutaneous coronary intervention
CABG : coronary artery bypass graft

한 결과는 Table 7과 같다. 주진단별로는 상세불명 부위의 급성 전층심근경색증(I213), 성별로는 여성, 연령별로는 75세 이상, 입원경로별로는 응급실을 통한 입원, 경피적 관상동맥 중재술(PCI)을 시행하지 않은 군, 폐혈증 등 동반상병을 동반한 군의 사망률이 높은 것으로 나타났다. 급성심근경색증 퇴원환자의 중증도 보정 사망에 영향을 미치는 동반상병에는 폐혈증, 직장 암, 기타 신경계통의 장애, 폐렴, 급성 신부전 및 만성 신부전 등이 있었다.

3.3.2 의사결정나무를 이용한 중증도 보정 사망 예측 모형 개발

3.3.2.1 모형 개발

Table 7. Severity-adjusted mortality rate model for acute stroke patients using logistic regression

Variables		Coeff.	Std. Err.	z-score	Odd Ratio	P> z
Constant		-3.672	0.429	-8.554		0.000
Principle Diagnosis	I211	-0.467	0.440	-1.060	0.627	0.289
	I212	-0.037	0.683	-0.055	0.963	0.956
	I213	1.074	0.379	2.835	2.926	0.005
	I214	-0.568	0.336	-1.693	0.566	0.090
	I219	0.673	0.294	2.286	1.960	0.022
Gender	Female	0.423	0.183	2.306	1.526	0.021
AGE	50-64	-0.113	0.382	-0.295	0.893	0.768
	65-74	0.961	0.359	2.678	2.614	0.007
	75+	1.398	0.360	3.880	4.046	0.000
Admission route	Ambulatory	-0.528	0.252	-2.093	0.590	0.036
PCI	YES	-0.567	0.177	-3.197	0.567	0.001
Septicemia(except in labor)	YES	2.191	0.742	2.951	8.942	0.003
Cancer of rectum and anus	YES	2.462	1.253	1.964	11.728	0.049
Coma: stupor; and brain damage	YES	3.014	0.736	4.095	20.366	0.000
Other nervous system disorders	YES	1.831	0.683	2.681	6.241	0.007
Peri-; endo-; and myocarditis; cardiomyopathy (except that caused by tuberculosis or sexually transmitted disease)	YES	1.266	0.474	2.673	3.546	0.008
Cardiac arrest and ventricular fibrillation	YES	2.285	0.398	5.736	9.823	0.000
Peripheral and visceral atherosclerosis	YES	1.239	0.563	2.202	3.453	0.028
Pneumonia (except that caused by tuberculosis or sexually transmitted disease)	YES	0.766	0.320	2.395	2.151	0.017
Gastritis and duodenitis	YES	1.163	0.544	2.139	3.198	0.032
Acute and unspecified renal failure	YES	1.057	0.355	2.973	2.877	0.003
Chronic renal failure	YES	0.770	0.369	2.089	2.159	0.037
Fracture of neck of femur (hip)	YES	2.558	0.918	2.788	12.914	0.005
Fracture of lower limb	YES	5.597	1.903	2.941	29.510	0.003
Shock	YES	2.267	0.428	5.295	9.648	0.000

Footnote) PCI : percutaneous coronary intervention

오픈소스 머신러닝 분석 도구인 KNIME software의 의사결정나무 기법으로 모형을 구축하였다. 데이터 전처

리, 모형 개발, 모형 평가 등 의사결정나무 기법을 이용한 급성심근경색증 퇴원환자의 동반상병 보정 모형 개발 워크플로어는 Fig. 2와 같다.

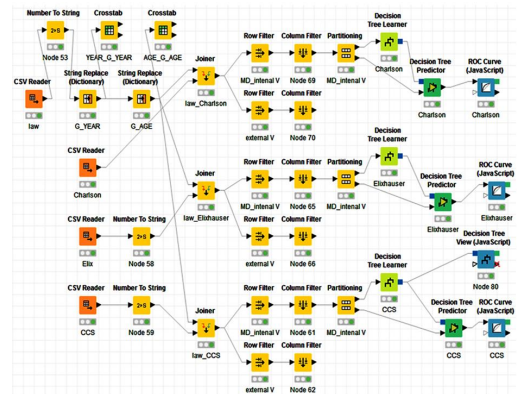


Fig. 2. KNIME Workflow for Decision Tree model development

3.3.2.2 모형 평가

의사결정나무를 이용한 모형 개발 및 평가 결과는 Table 8과 같다. 모델 1, 모델 2, 모델 3, 모델 4 모두에서 새로운 방법인 CCS 진단군 분류 변수를 이용한 동반상병 보정 방법이 포함된 모형의 내적검증 및 외적검증 AUC값이 가장 높은 것으로 나타났다. 4가지 모델 모형 중 모델 4의 CCS 진단군 분류 변수를 이용한 동반상병 보정 방법과 주진단, 성, 연령, 입원경로, 경피적 관상동맥 중재술(PCI) 시행 유무, 관상동맥 우회술(CABG) 시행 유무 모형의 내적검증 및 외적검증 AUC값이 각각 0.732, 0.697로 가장 높은 것으로 나타나 이를 이용하여 급성심근경색증 퇴원환자의 중증도 보정 사망 예측 모형을 개발하는 것이 가장 타당한 것으로 평가되었다.

Table 8. Decision tree model assessment using AUC

Variables		Internal Validation AUC	External Validation AUC
Model1	CCI	0.546	0.572
	ECI	0.596	0.576
	CCS index	0.667	0.646
Model2	CCI + Principle Diagnosis	0.543	0.573
	ECI + Principle Diagnosis	0.640	0.593
	CCS index + Principle Diagnosis	0.672	0.662

Model3	CCI + Principle Diagnosis + Gender + Age	0.699	0.686
	ECI + Principle Diagnosis + Gender + Age	0.686	0.676
	CCS index + Principle Diagnosis + Gender + Age	0.697	0.687
Model4	CCI + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.723	0.681
	ECI + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.720	0.704
	CCS index + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.732	0.697

Footnote) PCI : percutaneous coronary intervention
CABG : coronary artery bypass graft

3.3.2.3 모형의 결과

의사결정나무를 이용하여 CCS 진단군 분류 변수를 이용한 동반상병 보정 방법과 주진단, 성, 연령, 입원 경로, 경피적 관상동맥 중재술(PCI) 시행 유무, 관상동맥 우회술(CABG) 시행 유무를 독립변수로 한 급성심근경색증 퇴원환자의 중증도 보정 사망 예측 모형을 개발한 결과는 Fig. 3과 같다. 급성심근경색증 퇴원환자의 중증도 보정 사망에 영향을 요인은 연령, 주진단, 패혈증 유무, 심정지 및 심실세동 유무, 흡인성 폐렴 유무, 쇼크 유

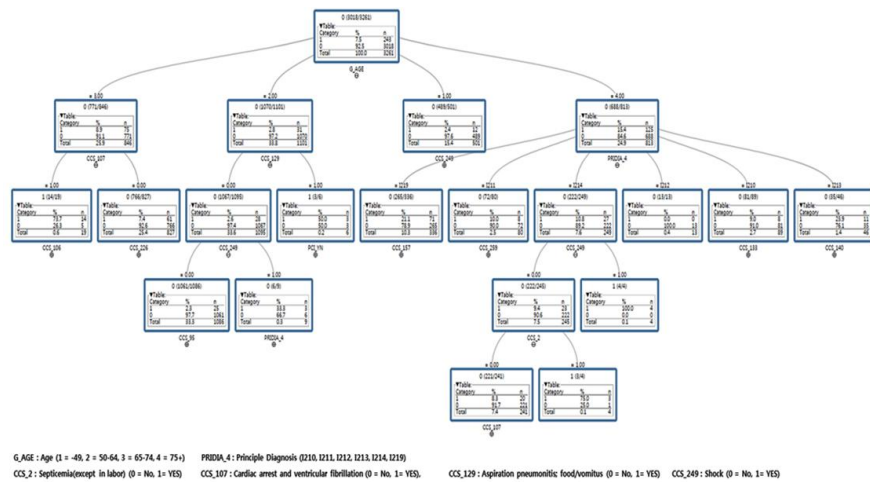


Fig. 3. Severity-adjusted mortality rate model for AMI patients using decision tree

무로 나타났다.

3.3.3 신경망을 이용한 중증도 보정 사망 예측 모형 개발

3.3.3.1 모형 개발

오픈소스 머신러닝 분석 도구인 KNIME software의 신경망 기법으로 모형을 구축하였다. 데이터 전처리, 모형 개발, 모형 평가 등 신경망 기법을 이용한 급성심근경색증 퇴원환자의 동반상병 보정 모형 개발 워크플로어는 Fig. 4와 같다.

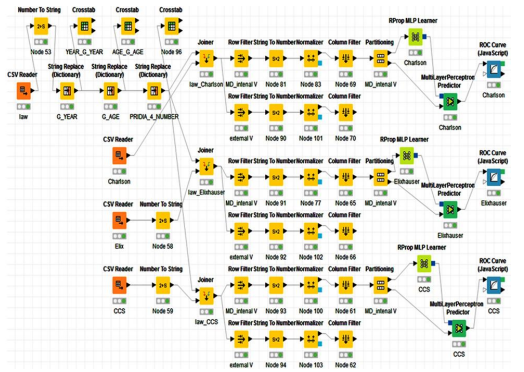


Fig. 4. KNIME workflow for neural network model development

3.3.3.2 모형 평가

신경망을 이용하여 급성심근경색증 퇴원환자의 동반상병 보정 모형을 개발하고, AUC 값을 이용하여 개발된

모형을 평가한 결과는 Table 9와 같다. 모델 1, 모델 2, 모델 3, 모델 4 모두 CCS 진단군 분류 변수를 이용한 동반상병 보정 방법이 포함된 모형의 내적검증 및 외적검증 AUC값이 가장 높은 것으로 나타났으며, 모델 1, 모델 2, 모델 3, 모델 4 모형 중 모델 4의 CCS 진단군 분류 변수를 이용한 동반상병 보정 방법과 주진단, 성, 연령, 입원경로, 경피적 관상동맥 중재술(PCI) 시행 유무, 관상동맥 우회술(CABG) 시행 유무 모형의 내적검증 및 외적검증 AUC값이 각각 0.724, 0.716으로 가장 높은 것으로 나타나 이를 이용하여 급성심근경색증 퇴원환자의 중증도 보정 사망 예측 모형을 개발하는 것이 가장 타당한 것으로 평가되었다.

Table 9. Neural network model assessment using AUC

Variables		Internal Validation AUC	External Validation AUC
Model1	CCI	0.521	0.468
	ECI	0.486	0.502
	CCS index	0.685	0.651
Model2	CCI + Principle Diagnosis	0.397	0.479
	ECI + Principle Diagnosis	0.439	0.458
	CCS index + Principle Diagnosis	0.698	0.658
Model3	CCI + Principle Diagnosis + Gender + Age	0.658	0.658
	ECI + Principle Diagnosis + Gender + Age	0.669	0.636
	CCS index + Principle Diagnosis + Gender + Age	0.695	0.673
Model4	CCI + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.675	0.650
	ECI + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.657	0.609
	CCS index + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.724	0.716

Footnote) PCI : percutaneous coronary intervention
CABG : coronary artery bypass graft

3.3.4 서포트 벡터 머신을 이용한 중증도 보정 사망 예측 모형 개발

3.3.4.1 모형 개발

오픈소스 머신러닝 분석 도구인 KNIME software의 서포트 벡터 머신 기법으로 모형을 구축하였다. 데이터 전처리, 모형 개발, 모형 평가 등 신경망 기법을 이용한 급성심근경색증 퇴원환자의 동반상병 보정 모형 개발 워크플로어는 Fig. 5와 같다.

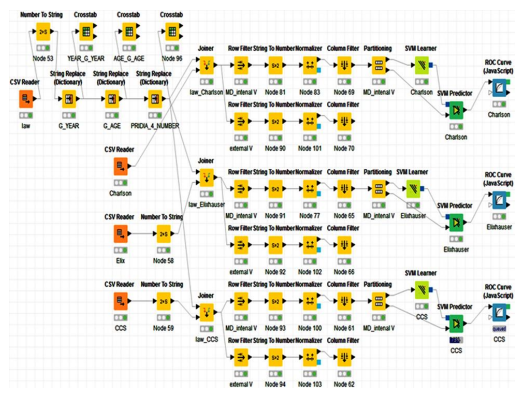


Fig. 5. KNIME workflow for support vector machine model development

3.3.4.2 모형 평가

서포트 벡터 머신을 이용하여 급성심근경색증 퇴원환자의 동반상병 보정 모형을 개발하고, AUC 값을 이용하여 개발된 모형을 평가한 결과는 Table 10과 같다. 모델 1, 모델 2, 모델 3, 모델 4 모두 Clinical classification software (CCS) 진단군 분류 변수를 이용한 동반상병 보정 방법이 포함된 모형의 내적검증 및 외적검증 AUC값이 가장 높은 것으로 나타났으며, 모델 1, 모델 2, 모델 3, 모델 4 모형 중 모델 4의 Clinical classification software(CCS) 진단군 분류 변수를 이용한 동반상병 보정 방법과 주진단, 성, 연령, 입원경로, 경피적 관상동맥 중재술(PCI) 시행 유무, 관상동맥 우회술(CABG) 시행 유무 모형의 내적검증 및 외적검증 AUC값이 각각 0.800, 0.748로 가장 높은 것으로 나타나 이를 이용하여 급성심근경색증 퇴원환자의 중증도 보정 사망 예측 모형을 개발하는 것이 가장 타당한 것으로 조사되었다.

Table 10. Support vector machine model assessment using AUC

Variables		Internal Validation AUC	External Validation AUC
Model1	CCI	0.593	0.604
	ECI	0.611	0.604
	CCS index	0.747	0.670
Model2	CCI + Principle Diagnosis	0.638	0.620
	ECI + Principle Diagnosis	0.663	0.629
	CCS index + Principle Diagnosis	0.757	0.686
Model3	CCI + Principle Diagnosis + Gender + Age	0.719	0.715
	ECI + Principle Diagnosis + Gender + Age	0.759	0.714
	CCS index + Principle Diagnosis + Gender + Age	0.796	0.729
Model4	CCI + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.740	0.741
	ECI + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.780	0.733
	CCS index + Principle Diagnosis + Gender + Age + Admission Route + PCI + CABG	0.800	0.748

Footnote) PCI : percutaneous coronary intervention
CABG : coronary artery bypass graft

4. 고찰

의료의 질평가는 크게 구조, 과정, 결과 평가로 분류할 수 있는데 최근에 결과 평가에 대한 관심이 크게 증가하고 있다. 결과 평가가 최근에 관심을 받는 이유는 적정비용으로 양질의 의료서비스를 제공할 필요성이 증대되기 때문이다. 특히, 결과 평가는 단순한 평가로 끝나지 않고 건강보험제도에서 수가에 반영을 하여 의료기관에 인센티브를 주는 방식을 취함에 따라 결과 평가가 급격히 확산되고 있다. 미국의 Center for Medicare & Medicaid Services(CMS)에서는 결과를 통하여 의료기관에 인센티브를 제공하고 있다. 우리나라의 건강보험심사평가원

에서는 적정성 평가에 결과 평가를 통하여 의료기관에 의료 질 향상 분담금을 차등 지급함에 따라 결과 평가가 매우 중요하다.

그러나 실제 결과 평가를 하기는 용이하지 않다. 왜냐하면 환자 마다 질환의 중증도가 다름에 따라 이를 보정하여 비교하는 것이 용이하지 않기 때문이다. 질환의 중증도 보정에 있어서 가장 중요한 요소가 동반질환이다. 동반질환을 보정하는 대표적 방법으로 CCI, ECI가 있으나[5,6] 이는 수작업 중심의 보정 방법이며, 동반질환 보정에 있어 제외되는 동반상병이 있음에 따라 정보화 시대의 맞춤형 보정방법으로는 제한점이 있다. 이에 본 연구에서는 맞춤형 보정방법을 개발하고 이의 타당성을 평가하기 위해 CCI, ECI처럼 주진단에 관계없이 제한된 동반질환의 가중치 및 동반질환 유무를 반영하여 동반질환을 보정하는 것이 아니라 퇴원환자의 모든 동반질환의 유무를 중증도 보정에 고려하여 주진단에 맞는 맞춤형 동반질환 보정 방법을 개발하고자 하였다. 이를 위해서 본 연구에서는 CCI, ECI와 다르게 퇴원환자의 동반질환 모두를 고려한 CCS 기준 259개 동반질환 중 주진단에 영향을 미치는 부진단을 선정하여 동반질환을 보정하고자 하였으며, 주진단에 영향을 미치는 부진단의 선정은 머신러닝 기법을 이용하는 것을 제안하였다. 본 연구에서 제안한 CCS 기준 동반질환 보정방법의 타당성을 평가하기 위해서 급성심근경색의 퇴원 시 사망에 대해 CCI, ECI와 본 연구에서 제안한 CCS 기준 동반질환 보정방법으로 퇴원환자의 중증도를 보정하고, 평가를 하였다. 평가를 위해서 질병관리본부의 퇴원손상자료를 수집하여 분석의 목적에 맞게 데이터를 정제하고 분석을 하였다. 모형의 타당성 평가는 내적 및 외적 타당도 평가를 실시하였다. 분석의 단계는 1단계에서는 동반질환만을, 2단계에서는 동반질환에 주진단을 추가하고, 3단계에서는 동반질환, 주진단에 성, 연령을 추가하고, 4단계에서는 여기에 입원경로, PCI여부, CABG여부를 추가하였다. 분석 방법은 각 단계별로 로지스틱회귀분석, 의사결정나무, 신경망, 서포트벡터머신을 이용하여 분석을 하였다. 모형의 평가지표는 AUC를 이용하였다.

모형평가 결과 CCS 기준 동반상병 보정방법의 모형 예측력이 가장 우수한 것으로 나타나 본 연구에서 제안한 CCS 기준 동반상병 보정방법으로 중증도를 보정하는 것이 타당한 것으로 나타났다. 또한, 모형의 개발은 머신러닝 기법인 서포트 벡터 머신이 가장 우수한 것으로 나

타나서 기존의 통계적인 분석방법인 로지스틱 회귀분석 보다는 서포트벡터 머신이 더 좋다는 것을 알 수 있다. 즉, 기존의 방법보다는 본 연구에서 제안한 맞춤형 보정 방법과 서포 벡터 머신 기법을 이용하는 것이 동반질환 보정이 더 우수하여 궁극적으로 신뢰할 수 있는 중증도 보정을 통해 결과평가에 큰 기여를 할 수 있다는 것을 확인할 수 있었다.

5. 결론

인구의 노령화, 신기술의 발달 등으로 의료비가 지속적으로 증가함에 따라 의료서비스의 효율화가 매우 필요한 시점이다. 이를 위해서는 의료기관의 의료서비스 결과에 대해서 결과평가를 하여 이를 기반으로 하여 인센티브제를 실시하여 의료기관이 결과향상을 위해서 노력하도록 유도할 필요가 있다. 이를 위해서 우선적으로 결과평가의 타당성을 보장할 수 있는 환자의 중증도 보정 방법이 개발되어야 한다. 기존의 중증도 보정방법인 CCI, ECI는 보정의 종류나 사망, 재원일수 등에 상관없이 동일한 방식으로 적용한다. 또한, 진단단의 특성을 반영하지 못하는 제한점이 있음에 따라 본 연구에서는 맞춤형 보정방법을 제시하고 이의 타당성을 평가하였다. 평가결과 본 연구에서 제안한 CCS 기준 동반상병 보정방법이 타당한 것이 나타났다. 그러나 본 연구의 결과는 단지 한 개 질환만을 대상으로 했고, 또한, 사망에 대한 중증도 보정 모형임에 따라 널리 활용하기에는 제한점이 많다. 따라서, 우선적으로 대상 질환을 확대할 필요가 있다. 또한, 사망이외의 재원일수 등 다른 지표에 대해서도 본 연구에서 제안한 방법의 타당성을 평가할 필요가 있다. 따라서 본 연구에서는 향후에 다음과 같은 노력이 이루어져야 한다는 것을 제안하고자 한다.

첫째, 본 연구에서는 우리나라 주요 사망 원인인 심근경색증 질환에 대한 보정방법을 제시하였으므로 향후 다른 질환에 대한 보정모형도 개발되어야 할 것이다.

둘째, 본 연구에서 개발한 방법을 사망이외의 재원일수 등의 보정방법으로 타당한 지에 대해서 연구가 필요하다.

셋째, 빅데이터를 이용하여 맞춤형 중증도 보정을 위해서는 정확한 질환 데이터가 수집되어야 한다. 따라서 중증도 보정방법의 개발 못지 않게 양질의 데이터 수집

이 중요하므로 양질의 데이터 수집을 위해 노력할 필요가 있다.

넷째, 맞춤형 중증도 보정방법이 널리 활용되기 위해서는 머신러닝을 이용하여 사망이나 재원일수 별로 또한 질환별로 각기 개발을 해야 한다. 따라서 맞춤형 보정방법이 활성화 되기 위해서는 이를 개발할 수 있는 인력양성이 매우 필요하다.

REFERENCES

- [1] J. H. Park, Y. M. Kim, S. S. Kim, W. J. Kim & S. H. Kang. (2012). Comparison of Hospital Standardized Mortality Ratio Using National Hospital Discharge Injury Data. *Journal of the Korea Academia-Industrial cooperation Society*, 13(4), 1739-1750. DOI : 10.5762/KAIS.2012.13.4.1739
- [2] J. H. Lim & J. Y. Park. (2011). The impact of comorbidity (the Charlson Comorbidity Index) on the health outcomes of patients with the acute myocardial infarction(AMI). *Korean J. of Health Policy & Administration*, 21(4), 541-564. DOI : 10.4332/KJHPA.2011.21.4.541
- [3] S. J. Kim, S. H. Kang, W. J. Kim & Y. M. Kim. (2011). The Variation Factors of Severity- Adjusted Length of Stay in CABG. *Journal of the Korean Society for Quality Management*, 39(3), 391-399. DOI : 10.7469/JKSQM.2011.39.3.391
- [4] T. K. Chung & S. H. Kang. (2013). The Comparison of Risk-adjusted Mortality Rate between Korea and United States. *The Journal of Digital Policy & Management*, 11(5), 371-384. DOI : 10.14400/JDPM.2013.11.5.371
- [5] S. O. Hong, Y. T. Kim, J. H. Park & S. H. Kang. (2015). The Variation of Factors of Severity- Adjusted Length of Stay (LOS) in Injury of Neck. *Health and Social Welfare Review*, 35(2), 561-583. DOI : 10.15709/hswr.2015.35.2.561
- [6] W. J. Kim, S. S. Kim, E. J. Kim & S. H. Kang. (2013). Severity-Adjusted LOS Model of AMI patients based on the Korean National Hospital Discharge in-depth Injury Survey Data. *Journal of the Korea Academia-Industrial cooperation Society*, 14(10), 4910-4918. DOI : 10.5762/KAIS.2013.14.10.4910
- [7] J. H. Lim & M. H. Nam. (2012). Development of Mortality Model of Severity-Adjustment Method of AMI

- Patients. *Journal of the Korea Academia-Industrial cooperation Society*, 13(6), 2672-2679.
DOI : 10.5762/KAIS.2012.13.6.2672
- [8] S. K. Jin & M. S. Bang. (2018). The Challenges of Public Policy Management for the 4th Industrial Revolution. *Journal of Digital Convergence*, 16(4), 39-47.
DOI : 10.14400/JDC.2018.16.4.039
- [9] K. Y. Lee & J. H. Kim. (2016). Artificial Intelligence Technology Trends and IBM Watson References in the Medical Field. *Korean Medical Education Review*, 18(2), 51-57.
DOI : 10.17496/kmer.2016.18.2.51
- [10] B. K. Choi, S. W. Ham, C. H. Kim, J. S. Seo, M. H. Park & S. H. Kang. (2018). Development of Predictive Model for Length of Stay(LOS) in Acute Stroke Patients using Artificial Intelligence. *Journal of Digital Convergence*, 16(1), 231-242.
DOI : 10.14400/kmer.2018.16.1.231
- [11] X. Cai, O. Perez-Concha, E. Coiera, F. Martin-Sanchez, R. Day, D. Roffe & B. Gallego (2016). Real-time prediction of mortality, readmission, and length of stay using electronic health record data. *J Am Med Inform Assoc*, 23, 553-561.
DOI : 10.1093/jamia/ocv110
- [12] A. Awada, M. Bader-El-Dena, J. McNicholasa & J. Briggsa. (2017). Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *International Journal of Medical Informatics*, 108, 185 - 195.
DOI : 10.1016/j.ijmedinf.2017.10.002
- [13] H. Nilsaz-Dezfouli, M. R. Abu-Bakar, J. Arasan, M. B. Adam & M. A. Pourhoseingholi. (2017). Improving Gastric Cancer Outcome Prediction Using Single Time-Point Artificial Neural Network Models. *Cancer Informatics*, 16, 1-11.
DOI: 10.1177/1176935116686062
- [14] Statistics Korea. (2017). *cause of death statistics 2016*. Statistics Korea Homepage.
http://kostat.go.kr/portal/korea/kor_nw/2/6/1/index.board?bmode=read&aSeq=363268
- [15] E. J. Han & J. S. Kim. (2015). Effects of Symptom Recognition and Health Behavior Compliance on Hospital Arrival Time in Patients with Acute Myocardial Infarction. *Korean Journal of Adult Nursing*, 27(1), 83-93.
DOI : 10.7475/kjan.2015.27.1.83
- [16] S. O. Hong, Y. T. Kim, Y. H. Choi, J. H. Park & S. H. Kang. (2015). Development of severity- adjusted length of stay in knee replacement surgery. *Journal of Digital Convergence*, 13(2), 215-225.
DOI : 10.14400/JDC.2015.13.2.215
- [17] H. H. Lee, S. H. Chung & E. J. Choi. (2016). A Case Study on Machine Learning Applications and Performance Improvement in Learning Algorithm. *Journal of Digital Convergence*, 14(2), 245-258.
DOI : 10.14400/JDC.2016.14.2.245
- [18] K. Y. Lee & J. H. Kim. (2016). Artificial Intelligence Technology Trends and IBM Watson References in the Medical Field. *Korean Medical Education Review*, 18(2), 51-57.
DOI : 10.17496/kmer.2016.18.2.51
- [19] M. A. Oh et. al. (2017). *A Study on Social security Big Data Analysis and Prediction Model based on Machine Learning*. Korea Institute for Health and Social Affairs Publishing.
<https://www.kihasa.re.kr/common/filedown.do?seq=39848>
- [20] H. H. Kim, S. B. Yang, Y. S. Kang, Y. B. Park & J. H. Kim. (2016). Machine Learning Approach to Blood Stasis Pattern Identification Based on Self-reported Symptoms. *Korean Journal of Acupuncture*, 33(3), 102-113.
DOI : 10.14406/acu.2016.011
- [21] I. S. Park, W. S. Yong, Y. M. Kim, S. H. Kang & J. T. Han. (2008). A Development of a Tailored Follow up Management Model Using the Data Mining Technique on Hypertension. *The Korean journal of applied statistics*, 21(4), 639-647.
DOI : 10.5351/kjas.2008.21.4.639
- [22] D. Y. Choi, K. M. Jeong & D. H. Lim. (2018). Breast Cancer Classification using Deep Learning-based Ensemble. *Journal of Health Informatics and Statistics*, 43(2), 140-147.
DOI : 10.21032/jhis.2018.43.2.140

백 설 경(Baek, Seol Kyung)

[정회원]



- 2000년 2월 : 아주대학교 공공정책 대학원 정책학과 보건의료전공
- 2009년 2월 : 아주대학교 의과대학 의료정보학과 박사과정 수료
- 1993년 11월 ~ 현재 : 아주대학교 병원 의무기록팀 팀장

· 관심분야 : 보건의료정보

· E-Mail : skbaek@aumc.ac.kr

박 혜 진(Park, Hye Jin) [정회원]



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

· 관심분야 : 보건의료정보, 공중보건, 건강증진
· E-Mail : hjpark@cu.ac.kr

강 성 홍(Kang, Sung Hong) [정회원]



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

· 관심분야 : 보건정보, 의무기록, 데이터마이닝, 건강증진
· E-Mail : hcmkang@hanmail.net

최 준 영(Choi, Joon Young) [정회원]



- 2007년 2월 : 원광대학교 정보관리학과(경영학석사)
- 2009년 2월 : 원광대학교 보건행정학과(보건행정학박사)
- 2012년 9월 ~ 현재 : 청암대학교 병원의료정보과 교수

· 관심분야 : 의료정보관리
· E-Mail : lemondote@naver.com

박 중 호(Park, Jong Ho) [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
· E-Mail : jh8283p@nate.com