

요인별 기관지천식에 대한 범주예측모형 개발

윤혜숙,¹ 나위진,¹ 최영진,¹ 김주화,¹ 오재원,¹ 김현희,² 장윤석,³ 유광하,⁴ 손건태⁵

¹한양대학교 의과대학 소아청소년과, ²가톨릭의과대학 소아청소년과, ³서울대학교 의과대학 알레르기내과, ⁴건국대학교 호흡기내과, ⁵부산대학교 통계학과

The development of patient-tailored asthma prediction model for the alarm system

Hey-Suk Yun,¹ Wee Jin Rah,¹ Young Jin Choi,¹ Joo-Hwa Kim,¹ Jae-Won Oh,¹ Hyun-Hee Kim,² Yoon-Seok Chang,³ Kwang-Ha Yoo,⁴ Keon-Tae Sohn⁵

¹Department of Pediatrics, Hanyang University College of Medicine, Seoul; ²Department of Pediatrics, Uijeongbu St. Mary's Hospital, The Catholic University of Korea College of Medicine, Uijeongbu; ³Division of Allergy and Clinical Immunology, Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seoul; ⁴School of Medicine, Konkuk University, Seoul; ⁵Department of Statistics, Pusan National University, Busan, Korea

Purpose: The increased incidence of asthma due to rising allergic diseases requires the prevention of worsening asthma. It is necessary to develop a patient-tailored asthma prediction model.

Methods: We developed causative factors for the asthma forecast system: infant and young children (0–2 years), preschool children (3–6 years), school children and adolescents (7–18 years), adults (19–64 years), old aged adult (> 64 years). We used the Emergency Department code data which charged the short-acting bronchodilator (Salbutamol sulfate) from Health Insurance Review and Assessment Service for the development of asthma prediction models. Three kinds of statistical models (multiple regression models, logistic regression models, and decision tree models) were applied to 40 study groups (4 seasons, 2 sex, and 5 age groups) separately.

Results: The 3 kinds of models were compared based on model assessment measures. Estimated logistic regression models or decision tree models were recommended as binary forecast models. To improve the predictability, a threshold was used to generate binary forecasts.


Conclusion: We suggest the binary forecast models as a patient-tailored asthma prediction system for this category. It may be needed the extended study duration and long-term data analysis for asthmatic patients for the further improvement of asthma prediction models. (*Allergy Asthma Respir Dis* 2016;4:328-339)

Keywords: Asthma, Asthma alarm system

서 론

기관지천식은 알레르기 등 과민 반응에 의한 만성 염증성 기도 질환으로써 모든 연령의 환자들에게서 기관지 과민반응성(airway hyperresponsiveness), 기도 협착(airway obstruction) 소견과 함께 기침, 천명, 호흡곤란 또는 흉부 압박감(chest tightness) 등의 증상을 보이는 질환이다. 천식은 현대사회에 급증하며 최근 전 세계적인 공중보건문제로 대두되고 있다.¹ 세계보건기구(World Health

Organization)의 보고에 따르면 전 세계적으로 천식의 유병률이 매 10년마다 50% 가량 증가하는 추세로, 전 세계적으로 천식 환자가 300만여 명, 천식으로 인한 사망자가 25만여 명에 달하는 것으로 추산되고 있으며, 천식은 소아에서 가장 흔한 만성 질환이다.^{2,3} 미국질병통제예방센터의 연구에 따르면 천식을 가진 소아와 청소년의 비율이 1980년의 3.6%에서 2003년의 5.8%까지 증가하였고,⁴ 우리나라도 천식 증상을 기준으로 봤을 때 소아 천식의 유병률은 1983년 5.7%, 1990년 10.1%였고, 2000년에는 초등생 13.0%, 중학생

Correspondence to: Jae-Won Oh  <http://orcid.org/0000-0003-2714-0065>
Department of Pediatrics, Hanyang University Guri Hospital, Hanyang University College of Medicine,
153 Gyeongchun-ro, Guri 11923, Korea

Tel: +82-31-560-2254, Fax: +82-31-552-9493, E-mail: jaewonoh@hanyang.ac.kr

*This study was supported by the grant of Korean Centers for Disease Control and Prevention (2011E3302300).

Received: March 20, 2016 Revised: July 27, 2016 Accepted: July 29, 2016

© 2016 The Korean Academy of Pediatric Allergy and Respiratory Disease
The Korean Academy of Asthma, Allergy and Clinical Immunology
This is an Open Access article distributed under the terms of the Creative
Commons Attribution Non-Commercial License
(<http://creativecommons.org/licenses/by-nc/4.0/>).

12.8%으로 증가하였다.^{5,7} 소아뿐만 아니라 50세 이후 성인에서도 천식의 유병률이 높아지고 있다. 특히 우리나라에서 2,467명을 대상으로 한 역학 연구에서 65세 이상에서 현증 천식의 유병률이 12.7%로 조사되어 세계에서 가장 빠르게 고령화 사회로 진입하고 있는 상황에서 사회 경제적으로 가장 중요한 질환이 될 수 있다.⁸ 기관지천식은 만성 질환으로 장기간 치료가 필요하기에 삶의 질을 떨어뜨리고 학습, 일 및 작업에 대한 능력을 떨어뜨리며, 치료비 부담, 생산의 감소로 인하여 개인 및 국가에 상당한 경제적 부담을 줄 수 있다.⁹ 미국의 경우, 천식으로 인한 직접적 의료비용 및 간접적 경제적 손실비용이 1994년 120억 달러 가량이 되었고 이는 그 10년 전과 비교하여 50% 이상 증가한 액수였는데, 주로 간접적인 경제적 손실비용으로 인한 것이었다.¹⁰ 이러한 간접비용은 환자 개인과 관련된 액수뿐만 아니라, 특히 환자가 소아라면 그 가족에 부담되는 비용을 포함한다. 우리나라의 경우에도 천식으로 인한 경제적 손실이 2004년에 총 20억 4천 달러에 달하였고(직접비용 46.9%, 간접비용 53.1%), 무형의 손실비용까지 합산하면 총 41억 1천 달러에 달하여 2004년의 국내총생산(gross domestic product)의 0.44%에 달하였다.¹¹ 따라서 기관지천식으로 인한 다방면의 손실을 최소화하기 위하여 국가적 차원에서의 공중보건 관리가 절실하다.

천식은 복합적인 원인의 질환으로, 유전적, 면역학적 이상과 환경적 원인으로 인하여 발병한다.¹² 천식 증상을 발현 및 악화시키는데 중요한 역할을 하는 환경적 유발 인자로는 공중항원(집먼지진드기, 동물의 털, 곰팡이, 꽃가루 등), 음식항원, 대기오염, 세균이나 바이러스 등의 미생물, 스트레스 등이 있다.¹³ 천식의 치료에서 이러한 위험 인자의 회피가 중요하며 세 단계로 구분할 수 있다. 1차 예방은 출생 전후 알레르겐에 노출될 가능성을 최소화시키기 위한 것이다. 그러나 영향을 줄 수 있는 요인은 제대 기간에 따라 다르고 매우 복잡하여 산전 알레르기 위험 인자를 알아내어 회피하는 것은 쉽지 않다. 위험 인자가 있는 어머니에서 태어난 아이의 경우 출생 전후에 알레르겐에 노출되는 것을 피함으로써 출생 후 발생할 수 있는 알레르기 면역반응을 변화시킬 수 있다. 2차 예방은 알레르겐에 감작된 후 다시 노출되는 것을 회피함으로써 천식 발병을 억제하는 것으로 영유아시기부터 초기에 시작되어야 한다. 3차 예방은 천식 환자에서 증상 악화는 알레르겐, 공해, 식품, 약물 등 다양한 자극에 의해 유발되므로 환자의 위험 인자를 파악하고 이에 노출되는 것을 방지하는 것이 천식을 조절할 수 있고 약물 투여량을 줄이는 데 중요한 방법이다.¹⁴

천식 치료 및 예방을 위해 이러한 환경적 요인들과 천식과의 연관성에 대한 연구는 활발히 이루어지고 있으나, 인구적인 차원에서 (population level) 환경적 요인들의 실제적인 위험 수준을 조사하고 천식 위험성에 대한 조기 경고 및 대비할 수 있는 시스템은 아직 미비하다.¹⁵ 천식 발현 또는 악화에 영향을 주는 기상 요소와 대기오염의 변화 양상에 따라 질환의 발병 여부 및 중증도가 어떻게 달라지

는지를 관찰하는 것은, 지속적으로 변화하는 환경적 요인과 관련하여 천식 환자들이 어떻게 영향을 받을지 예측하는 데 유용하다.¹⁶

우리나라 기상청에서는 2005년도부터 건강보험 청구 자료를 기반으로 산출한 천식 입원율을 기상정보와 연계하여 9개 시도별로 천식가능지수를 예보하고 있고, 2009년도부터 질병관리본부에서도 응급콜센터 자료를 기반으로 서울지역을 대상으로 천식예보제를 시행하고 있으나 그 예보제의 기반이 되는 천식 입원율과 응급콜센터 자료의 실효성 및 천식예보의 정확성에 문제가 제기되고 있다. 기존에 시행하고 있던 질병관리본부의 천식예보제와 건강보험 청구자료와 기상정보를 바탕으로 한 천식예보제보다 좀 더 구체적이고 환자맞춤형인 천식예보모델을 개발하고자 본 연구를 시작하게 되었으며 본 연구는 국내에서 지속적으로 관측되어온 여러 가지 예측 인자(기상 요소, 대기오염, 환경 인자, 보건 인자)를 사용하여, 성별, 연령, 계절에 따라 천식 악화 양상이 어떻게 달라지는지를 보여주는 실용성 있는 천식예보모형을 개발하고자 하였다.

대상 및 방법

1. 환자 자료

본 연구에서는 서울지역에서 응급실을 방문하여 천식 진단하여 치료받은 환자를 대상으로 2006년 7월 1일부터 2010년 12월 31일까지의 기간에 해당하는 자료를 수집하였다. 환자 수를 전수조사 또는 표본조사하기가 현실적으로 어려우므로, 서울지역 응급실 외래 환자에 대한 심사평가원자료를 대체자료(proxy data)로 사용하였다. 심사평가원에서 제공하는 약제 및 환자정보자료는 건강 보험자를 대상으로 서울지역 요양기관에서 심사평가원에 청구한 자료이다. 약제정보는 응급의료 관리료 수가코드로 청구한 건들 중에서 Salbutamol sulfate (에어로솔제, 분무제 및 흡입제) 약품을 청구한 수진자수이다. 환자 정보자료는 다음과 같이 구분하여 예측모형 개발에 사용하였다. (1) 계절(봄, 여름, 가을, 겨울), 성별(남자, 여자), 연령군(5개 연령군)으로 구분하여 총 40 유형(4계절×2성별×5연령군)으로 분류하였다. (2) 계절을 나누는 기준은 꽃가루와 인플루엔자 유행 시기를 고려하여, 봄철은 3월 11일-6월 10일, 여름철은 6월 11일-9월 10일, 가을철은 9월 11일-11월 30일, 겨울철은 12월 1일-3월 10일로 구분하였다. (3) 연령군은 영아기(0-2세), 학령전기(3-6세), 학령기(7-18세), 청장년기(19-64세), 노년기(65세 이상)로 구분하였다.

2. 범주 구분

범주는 2범주(지속관리, 주의)로 구분하였으며, 범주유형별로 각각 예측모형을 개발하여 예측 결과를 비교한 후 천식예보제에 활용하기로 하였다. 범주 분류 기준은 계절별, 성별, 연령군별로 구분된 심사평가원 자료 분포의 제75백분위수를 사용하였다. 환자

수가 제75백분위수 이하이면 ‘지속관리’, 제75백분위수보다 크면 ‘주의’로 구분하였다. 예측모형 개발에서 ‘지속관리’는 0, ‘주의’는 1로 표시하였다.

3. 예측 인자

1) 천식 유발 인자

천식발작의 유발 인자로는 유전적 요인, 운동, 감기 바이러스, 대기오염, 약물, 음식물, 기온, 습도, 동물 털, 곰팡이, 꽃가루, 집먼지진드기, 식도역류, 스트레스 등 많은 요인이 있다. 예측모형 개발을 위하여 수집할 수 있는 자료들 중에서 질병관리본부와 천식 전문가의 의견을 바탕으로 기상자료, 대기오염자료, 알레르기꽃가루 자료, 보건자료들로 이루어진 인자(Table 1)를 사용하였고 모두 일자료(daily data)로 구성하였다. 기상자료는 기온, 기압, 습도, 일조량이며, 대기오염자료는 지상오존과 황사자료(황사관측 유무, 황사농도)이며, 보건자료는 병원 환자 중 독감 환자 비율을 사용하였다.

2) 지속성 인자와 요일 인자

심사평가원 자료의 자기상관성(autocorrelation)을 조사하여 유의한 시차상관(lag correlation)이 있는 과거자료를 지속성 인자로 정의하며 예측 인자에 포함을 시켜 예측모형 개발에 사용하였다.

천식발작은 생활환경과 관련이 있으므로 생활 패턴의 중심이 되는 요일 효과를 예측 인자에 포함시켰다. 요일 인자는 계량치변수가 아니므로 가변수(dummy variable)를 사용하여 D1(일요일), D2(월요일), D3(화요일), D4(수요일), D5(목요일), D6(금요일)로 구분하였다. 해당 일에 1의 값을 배당하고, 아니면 0을 배당한다. 월요일은 (D1, D2, D3, D4, D5, D6) = (0, 1, 0, 0, 0, 0)이 되며, 토요일은 (D1, D2, D3, D4, D5, D6) = (0, 0, 0, 0, 0, 0)이 된다.

3) 잠재적 예측 인자

심사평가원의 환자자료의 자기상관성(autocorrelation)을 알아 보기 위하여 시차상관분석(lag correlation)을 수행한 결과 모든 유형에서 7일 전 자료가 유의한 자기상관이 있는 것으로 나타났다. 심사평가원의 환자자료와 천식 유발 인자들 사이의 교차상관(cross correlation)을 수행하여 유의한 시차를 선택하여 잠재적 예측 인자(potential predictors)를 선정하여 Table 1에 정리하였다.

Table 1에서 변수 표기는 유의시차가 1인 경우 변수 앞에 L을 붙이고, 유의시차가 m인 경우는 Lm을 붙였다. 예를 들어, Table 1에서 남자영아기(M1) 지속성 인자는 L7M1으로 표기하였다. L은 lag을 의미하고 7은 시차이다. L7W5는 여자노년기(W5)의 지속성 인자를 뜻한다. 남자 영아기(M1) 경우에 일평균기온(T)은 유의시차가 1이므로 LT로 표기하고, 지상오존농도(O3)는 유의시차가 4이므로 L4O3로 표기하였다. 다른 변수도 동일한 규칙으로 표기하였다.

4. 천식 예보모형 개발을 위해 적용된 통계모형

본 연구는 익일예보를 대상으로 하였다. 즉, 오늘까지 예측 인자들을 사용하여 내일의 천식 환자의 수에 대한 범주예측을 생산하는 모형개발을 목적으로 하였다. 심사평가원 자료는 반응변수로 사용하고, Table 1의 잠재적 예측 인자들을 예측 인자로 사용하였다. 2010년 자료를 모형검증자료로 사용하였다. 모형개발 결과는 예측 결과를 2×2 분할표로 요약하여 예측성 평가를 실시하고 비교하였다.

본 연구에서는 세 가지 통계모형을 적용하였다. 중회귀모형을 통해 천식 환자 발생 수의 계량치 예측 값을 생성한 후 범주 분류 기준에 따라 범주예측치를 생성하였다. 순서형 로지스틱 회귀모형과 의사결정나무모형에서는 확률예측치를 생산하며, 최대확률 또는

Table 1. Potential predictors of asthma

Predictor	Interval	Type	M1	M2	M3	M4	M5	W1	W2	W3	W4	W5
A	Daily	Quantitative	7	7	7	7	7	7	7	7	7	7
T	Daily	Quantitative	1	1	1	1	1	1	1	1	1	1
DT	Daily	Quantitative	1	1	1	1	1	1	1	1	1	1
MH	Daily	Quantitative	1	1	1	1	1	1	1	1	1	1
PR	Daily	Quantitative	1	1	1	1	1	1	1	1	1	1
HS	Daily	Quantitative	1	1	1	1	1	1	1	1	1	1
OZ	Daily	Quantitative	4	1	1	1	1	3	1	1	2	1
PL	Daily	Quantitative	1	1	1	1	1	1	1	1	1	1
FL	Weekly	Quantitative	1	1	1	1	1	1	1	1	1	1
PM	Daily	Quantitative	1	1	6	6	1	1	1	7	1	1
YS	Daily	Binary	1	2	2	3	1	1	2	6	1	2
D1-D6	Daily	Dummy	-	-	-	-	-	-	-	-	-	-

Numbers indicate the significant lag times (days) between environmental factors and the Health Insurance Review and Assessment Service data (occurrence of asthma symptoms).

A, autocorrelated factor; T, mean temperature; DT, daily range; MH, minimum humidity; PR, pressure; HS, hours of sunshine; OZ, concentration of ozone; PL, pollen; FL, proportion of flu patients; PM, concentration of yellow sand; YS, presence of yellow sand; D, day of the week; M, man; W, woman.

문턱치를 고려하여 범주예측치를 생성하였다. 또한, 범주예측모형 평가와 최적 문턱치를 결정하기 위하여 예측성 평가측도(skill score)를 고려한다. 계절별, 성별, 연령군별로 구분된 자료 각각에 대하여 세 모형을 적용하여 얻은 범주예측치를 비교하고, 예측성 평가측도인 정분류율(hit rate, HR), 탐지확률(probability of detection, POD), 허위경고율(false alarm rate, FAR)을 구하여 비교분석하여 최적 예측모형을 제안하였다.

1) 중회귀모형

잠재적 예측 인자를 설명변수로 하는 중회귀모형식은 다음과 같다. 변수선택은 단계별 회귀방법(stepwise regression method)을 적용하였다.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \epsilon$$

중회귀모형에서 생성된 계량치 예측치는 심사평가원 자료에 대한 예측치이므로 이 값을 그대로 예보치로 사용할 수는 없다. 계량치는 음수가 나올 수 있으며, 이는 환자 수가 0인 강한 증거로 해석하여 0으로 추정한다. 유형별로 결정된 범주 분류기준에 따라 범주 예측치로 변환시켜 범주 예측에 사용한다.

2) 로지스틱 회귀모형

심사평가원 자료와 범주 분류기준에 따라 생성된 관측범주와 이차 잠재적 예측 인자를 이용하여 로지스틱 회귀모형(logistic regression model)을 적용한다. 반응변수 Y가 범주 값(0, 1)을 가질 때, 이 범주 예보를 결정하는 누적확률식과 알고리즘은 다음과 같다.

$$P(\text{주의}) = P(Y=1) = \frac{\exp(b_0 + \beta' X)}{1 + \exp(b_0 + \beta' X)}$$

$$P(\text{지속관리}) = 1 - P(\text{주의})$$

위 수식에서 $X = (X_1, X_2, \dots, X_p)$ 는 설명변수로 이루어진 열 벡터이다. 본 연구에서는 이 단계로 범주예측치를 생성하였다. 첫 단계로, 예측 인자값이 주어진 경우 추정된 위 수식으로 $P(Y=1)$ 의 추정치를 구한다. 둘째 단계로, 문턱치(threshold)를 고려하여 $P(Y=1)$ 의 추정치가 문턱치보다 크면 '주의'로, 아니면 '지속관리'로 범주예보치를 생성한다. 최적문턱치는 다음의 절차로 결정한다. 문턱치를 0부터 1까지 변화시켜가며, 2범주 예측치들을 생성하고, 예측성 평가를 수행하여 예측성이 가장 우수한 문턱치를 선택한다.

3) 의사결정나무모형

관측범주와 예측 인자를 이용하여 의사결정나무모형을 적용하였다. 반응변수가 이산형인 경우에는 반응변수의 각 범주에 속하는 빈도에 기초하여 분리가 일어나며, 본 연구에서는 엔트로피 지수(entropy index)를 분리 기준으로 사용하였다. 의사결정나무모형은 분류 기준을 통해 예측치에 대한 범주를 구분할 수 있으며, 모

Table 2. Assessment of model validity and predictability

Actual case	Forecasted category		
	Continuous management	Attention	Total
Continuous management	A	B	A+B
Attention	C	D	C+D
Total	A+C	B+D	A+B+C+D

A, negative correction; B, false alarm; C, miss; D, hit.

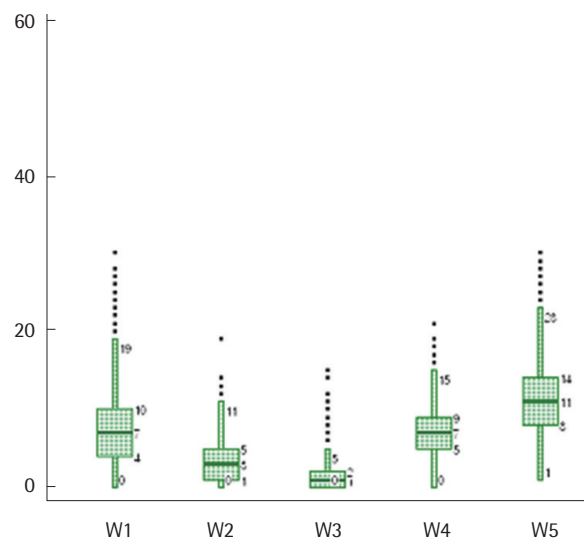
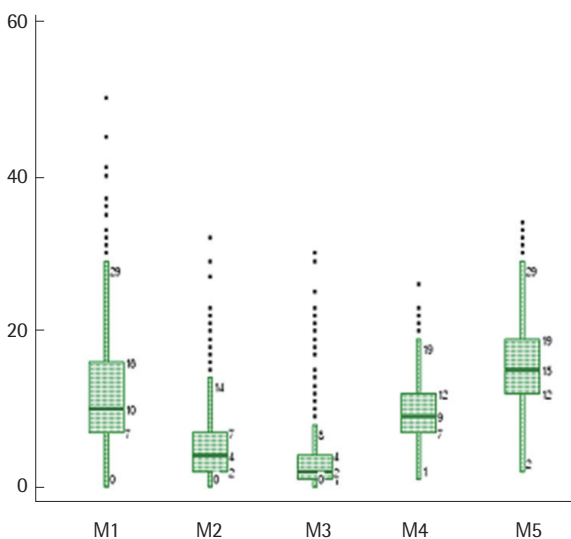


Fig. 1. Box plot showing distribution of patients with asthma by sex and age group. M, man subject; W, woman subject. M1 and W1: 0-2 years old, M2 and W2: 3-6 years old, M3 and W3: 7-18 years old, M4 and W4: 19-64 years old, M5 and W5: 65 years old.

형적합을 통하여 확률예측치를 생성한다. 의사결정나무모형에서도 로지스틱회귀모형과 동일한 과정으로 문턱치를 사용하여 범주 예측치를 생성하였다.

4) 예측성 평가 척도와 문턱치

모형의 예측성을 비교하기 위하여 범주 예측 결과를 2×2 분할 표로 요약하여 예측성 평가척도들(skill scores)을 계산하여 비교하였다. 이 범주의 경우 HR, POD, FAR을 사용하였다. 각 예측성 평가 척도는 Table 2의 기호를 사용하여 다음 식으로 표현된다.

$$HR = (A+D)/(A+B+C+D)$$

$$POD = (D)/(C+D)$$

$$FAR = B/(B+D)$$

HR은 실제 상태를 정확히 예측한 비율을 나타낸다. POD는 실제가 주의일 경우 정확히 분류한 비율이다. FAR은 실제가 지속 관리인데 주의라고 잘못 예측한 경우이다. 주의인 경우 천식 환자가 많이 발생한다는 것을 뜻하므로, 지속 관리보다 주의에 더 비중을 두게 된다. 따라서 실제가 주의일 때 정확히 예측한 것을 양의 정분류(hit), 실제가 지속관리일 때 정확히 예측한 것을 음의 정분류(negative correction)로 구분하여 정의한다. HR이 높더라도 POD가 낮으면, 예측모형으로 활용할 수 없다. HR과 POD는 높을수록 좋고, FAR은 낮을수록 좋은 모형이 된다. 본 연구에서는 탐지확률이 70% 이상, 허위경고율이 50% 미만을 기준으로 최적 문턱치와

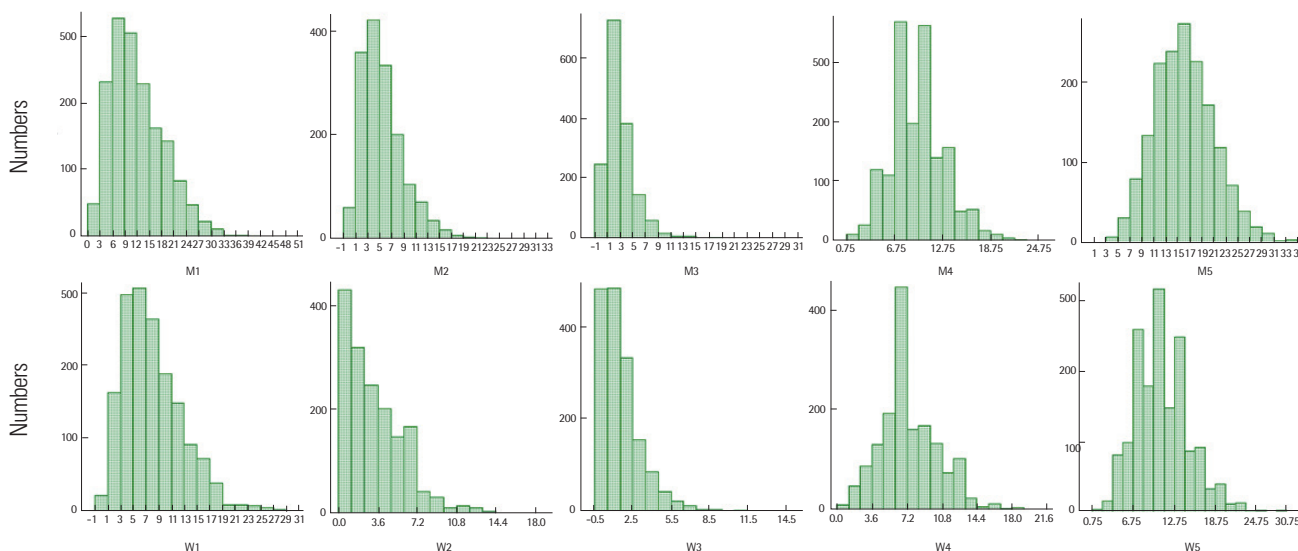


Fig. 2. Histogram showing distribution of patients with asthma by sex and age group. M, man subject; W, woman subject. M1 and W1: 0–2 years old, M2 and W2: 3–6 years old, M3 and W3: 7–18 years old, M4 and W4: 19–64 years old, M5 and W5: 65 years old.

Table 3. Asthma cases in the HIRA dataset according to sex, age group (1–5), and season

Group	Spring		Summer		Autumn		Winter	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
M1	14.0	6.9	7.6	5.3	13.7	6.5	12.4	6.2
M2	5.5	3.3	3.2	2.6	7.5	4.5	4.5	3.0
M3	2.7	2.0	1.6	1.8	4.7	4.5	2.1	1.9
M4	10.2	3.8	9.0	3.0	9.9	3.4	9.7	3.5
M5	17.5	5.0	14.1	5.0	14.7	4.7	16.2	4.9
W1	9.1	5.1	4.8	3.6	9.1	4.8	7.8	4.2
W2	3.8	2.7	2.1	1.7	4.7	3.0	2.9	2.1
W3	1.6	1.6	0.9	1.2	2.6	2.2	1.2	1.4
W4	7.7	3.0	6.1	2.6	7.8	3.4	8.0	3.2
W5	12.3	4.0	9.5	3.5	10.4	3.6	12.2	4.4
Total	84.5	21.0	58.9	16.2	85.1	21.0	76.9	17.3

HIRA, Health Insurance Review and Assessment Service; SD, standard deviation; M, man; W, woman.

이 범주예측모형을 선택하였다.

결 과

1. 환자의 분포 특성

심사평가원 환자 자료에서 성별, 연령 군별 기초 통계치는 요약하여 분포의 개관을 상자그림(box plots)으로 표현하였다. 성별 연령 군별로 구분하여 환자 수를 비교하면 남자와 여자 모두 노년기, 영아기, 청장년기, 학령전기, 학령기 순으로 나타났다(Fig. 1). 성별 연령 군별 분포를 보면 영아기, 학령전기, 학령기는 왼쪽으로 치우친 비대칭적 분포 형태를 보이고 있으며, 청장년기와 노년기는 정규 분포에 가까운 대칭적 분포를 보이고 있다(Fig. 2). 각 연령군에서 여자의 분포는 남자의 분포와 형태가 비슷하지만, 여자 환자 수가 남자보다 15%~30% 적게 나타났다. 연령군별로 환자비율을 계산하면, 영아기는 73%, 학령전기는 82%, 학령기는 69%, 청장년기는 87%, 노년기는 77%로 나타났다. 계절별 성별 연령 군별 기초 통계치에서 계절별로 구분하면 연령군별로 차이가 있으나 전체적으로는 가을철, 봄철, 겨울철, 여름철 순으로 나타났으며, 가을철과 봄철은 비슷한 환자 수를 보이고 있다. 영아기 남자와 여자는 모두 봄철, 가을철, 겨울철, 여름철, 학령전기와 학령기 남자와 여자는 모두 가을철, 봄철, 겨울철, 여름철, 청장년기 남자는 봄철, 가을철, 겨울철, 여름철, 청장년기 여자는 겨울철, 가을철, 봄철, 여름철, 노년기 남자와 여자는 모두 봄철, 겨울철, 가을철, 여름철 순으로 나타났다(Table 3). 이들을 종합하여(4계절)×(남자, 여자)×(5개 연령군)의 40 종류에 대한 천식 환자 집단을 구분하여 예측모형을 개발하였다.

2. 환자 자료와 예측 인자 사이의 상관분석

예측모형 개발을 위하여 상관성이 없는 예측 인자를 제거하기 위하여 심사평가원 환자 자료와 예측 인자 사이의 상관분석을 수행하여 잠재적 예측 인자를 선정하였다(Table 4). 대부분의 경우 지속성 인자, 일 평균 기온, 꽃가루 자료, 독감 환자 비율이 선정되었다. 영아기 여자의 경우 봄과 여름에는 일교차가 유의하며, 가을과 겨울에 최저습도가 유의하게 나타났다. 노년기 경우 봄과 여름에는 지상오존농도가 유의하며, 가을과 겨울에는 최저습도와 기압이 유의하게 나타났다.

3. 2 범주 예측모형 개발

이 범주 예측모형의 최종 제안 모형은 정분류율, 탐지확률, 허위 경고율을 참고하여, 정분류율과 탐지확률을 높이고 허위경고율을 낮추는 문턱치를 선정하여 범주예측에 사용하였다. 모형 선택의 기준은 정분류율과 탐지확률은 70% 이상, 허위경고율은 50% 미만으로 정하였고, 세 모형이 모두 기준에 못 미치는 경우에는 예측성

Table 4. Potential predictors

Predictors	Spring					Summer					Autumn					Winter														
	M1	M2	M3	M4	M5	W1	W2	W3	W4	W5	M1	M2	M3	M4	M5	W1	W2	W3	W4	W5	M1	M2	M3	M4	M5	W1	W2	W3	W4	W5
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OZ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

A circle means that the predictor is significantly correlated with the corresponding group. A, autocorrelated factor; T, mean temperature; DT, daily range; MH, minimum humidity; PR, pressure; HS, hours of sunshine; OZ, concentration of ozone; YS, presence of yellow sand; PM, concentration of yellow sand; PL, pollen; FL, proportion of flu patients; M, man; W, woman.

Table 5. Comparison of binary forecasting models for the spring season

Group	Model	Threshold	Skill scores (%)						Final predictors
			Whole data			Validation			
			HR	POD	FAR	HR	POD	FAR	
M1	REG	*	79.78	47.87	34.78	47.83	32.26	23.08	A T FL D1
	LOGISTIC	0.25	74.24	73.40	49.64	58.70	54.84	22.73	A FL PN
	TREE	0.50	81.16	54.26	32.90	42.39	22.58	26.32	A
M2	REG	*	77.01	47.71	33.33	55.43	30.36	10.53	A PR FL D1 D5 D6
	LOGISTIC	0.25	70.01	72.48	49.68	63.04	62.50	27.08	A PR FL D1
	TREE	0.20	76.62	61.47	40.71	39.13	00.00	**	A PR FL D1
M3	REG	*	77.29	26.92	17.65	59.78	30.23	35.00	PR FL PN D1 D4
	LOGISTIC	0.25	70.11	72.38	51.59	67.39	72.09	36.73	PR FL PN D1
	TREE	0.25	67.12	75.24	54.60	59.78	37.21	38.46	A T FL PN D1
M4	REG	*	70.65	13.79	27.27	61.95	34.29	50.00	FL D2
	LOGISTIC	0.40	69.02	28.45	48.44	65.22	22.86	38.46	FL D2
	TREE	0.50	72.83	18.10	19.23	65.22	22.86	38.46	FL D2
M5	REG	*	74.79	27.17	48.98	55.43	37.50	51.61	A FL PN D1 D2
	LOGISTIC	0.30	71.75	53.26	54.21	50.00	72.09	36.73	A FL PN D2
	TREE	0.50	78.67	41.30	37.70	46.74	32.50	62.86	A FL PN
W1	REG	*	78.12	41.49	38.10	50.00	19.64	8.33	A HS FL D1
	LOGISTIC	0.25	75.62	74.47	47.76	67.39	60.71	19.05	A HS
	TREE	0.25	84.21	70.21	30.53	52.17	37.50	30.00	A T DT HS PM D6
W2	REG	*	73.13	35.54	30.65	54.35	39.34	17.24	A OZ FL D1
	LOGISTIC	0.25	62.60	74.38	53.61	75.00	85.25	21.21	A FL
	TREE	0.50	79.50	45.45	12.70	56.52	36.07	4.35	A T DT PM FL D6
W3	REG	*	66.30	40.24	28.26	64.13	54.10	13.16	PR FL D1 D2
	LOGISTIC	0.35	58.42	74.39	47.64	67.39	77.05	25.40	PR FL D1 D2
	TREE	0.40	72.02	70.81	32.14	61.96	86.89	33.75	A T MINH PR FL
W4	REG	*	72.55	0.99	50.00	68.48	6.67	33.33	DT FL D1 D2
	LOGISTIC	0.40	59.78	60.40	63.91	52.17	80.00	61.29	DT FL D2
	TREE	0.25	71.74	72.28	51.01	61.96	60.00	56.10	DT MH FL PN D1 D2
W5	REG	*	68.98	22.95	39.13	52.17	47.06	61.90	A OZ FL D2
	LOGISTIC	0.35	67.87	49.18	47.37	60.87	38.24	53.57	A D2
	TREE	0.30	73.68	46.72	34.48	63.04	0.00	0.00	A FL PN D2 D3

A threshold cannot be used to determine the category of symptoms in the multiple regression model.

HR, hit rate; POD, probability of detection; FAR, false alarm rate; A, autocorrelated factor; T, mean temperature; DT, daily range; MH, minimum humidity; PR, pressure; HS, hours of sunshine; OZ, concentration of ozone; PL, pollen; FL, proportion of flu patients; PM, concentration of yellow sand; YS, presence of yellow sand; D, day of the week; M, man; W, woman.

*No numerical value. **There is no significant predictive factor or dummy variable denoting day of week.

이 가장 좋은 모형으로 선택하였다. 확률예측치를 통해 범주예측치를 정하는 로지스틱회귀모형과 의사결정나무모형은 예측성 평가 측도 향상을 위하여 문턱치(threshold)를 사용하여 범주예측치를 정하였다. (1) 범주 분류 기준으로 문턱치를 사용하지 못하는 중 회귀모형의 예측성 평가 측도는 정분류율을 제외하고는 저조하게 나왔다. (2) 로지스틱회귀모형과 의사결정나무모형은 확률예측치를 생산한다. 이 범주의 경우 HR과 POD를 70% 이상으로 크게 하고, FAR을 50% 이하로 작게 하는 예측규칙이 바람직한 것으로 알려져 있다. 이를 위하여 범주예측치를 결정할 때 최대 확률을 기준

으로 분류하는 것보다 문턱치 기준으로 분류하는 것이 더 우수한 예측성을 보인다. 즉, 확률예측치가 과소 추정되거나 과대 추정되는 경우에는 0.5 기준의 분류는 좋은 예측 결과를 보이지 않으므로, 확률예측치를 그대로 예보에 사용할 수 없다. 이런 경우, 특히 지속 관리가 75%, 주의가 25%로 구성된 이 범주의 경우에는 문턱치를 고려한 범주예측치 생성이 바람직하다. 문턱치를 사용하는 경우 Pr(주의)가 문턱치보다 크면 ‘주의’라고 예측을 하고, 아니면 ‘지속 관리’로 예측한다. 최적 문턱치를 찾기 위하여 문턱치를 0.01부터 0.99까지 0.01단위로 변경시켜 범주예측치를 생성하고 해당 2

Table 6. Comparison of the binary forecasting models for the summer season

Group	Model	Threshold	Skill scores (%)						Final predictors
			Whole data			Validation			
			HR	POD	FAR	HR	POD	FAR	
M1	REG	*	78.75	53.17	33.00	63.04	51.52	5.56	A T L4OZ FL D1
	LOGISTIC	0.25	72.29	71.43	48.28	67.39	78.79	23.53	A FL D1
	TREE	0.25	72.52	72.22	48.00	42.39	19.70	00.00	A FL D1
M2	REG	*	70.90	38.41	36.26	61.96	40.82	23.08	A PR D1
	LOGISTIC	0.30	65.36	68.21	49.76	58.70	63.27	39.22	A PR D1
	TREE	0.25	65.82	84.11	49.40	56.52	22.45	15.38	A FL PR T D1
M3	REG	*	57.97	23.00	37.84	55.43	17.78	33.33	A T MH D1
	LOGISTIC	0.40	52.96	78.43	50.46	57.61	64.44	44.23	T MH D1
	TREE	0.50	56.14	18.14	41.27	48.91	11.11	58.33	D1
M4	REG	*	69.09	00.00	**	59.78	00.00	**	D2
	LOGISTIC	0.20	30.91	100.0	69.09	40.22	100.0	59.78	D2
	TREE	0.30	59.82	61.03	59.31	66.30	24.32	25.00	A PL D2
M5	REG	*	73.67	24.81	34.69	44.57	14.04	20.00	A OZ FL D1 D2
	LOGISTIC	0.25	65.36	74.42	54.93	60.87	56.14	25.58	A FL D2
	TREE	0.35	68.13	73.64	52.26	45.65	17.54	23.08	A OZ FL D2
W1	REG	*	75.75	44.44	38.46	36.96	10.77	0.00	A T OZ FL D1
	LOGISTIC	0.25	69.98	73.02	51.06	41.30	23.08	21.05	A T FL D1
	TREE	0.30	84.53	71.43	25.62	50.00	38.46	19.35	A T OZ FL D1 D3 D4
W2	REG	*	69.52	16.43	39.47	46.47	5.88	25.00	A T PR FL D1
	LOGISTIC	0.30	62.60	74.38	53.61	54.35	21.57	15.38	FL D1
	TREE	0.30	66.74	77.86	50.90	47.83	41.18	46.15	A T PR FL D4 D5
W3	REG	*	56.14	46.88	32.20	47.83	33.33	28.57	T MH PR D1 D2
	LOGISTIC	0.45	60.82	92.97	39.29	66.30	95.00	32.94	PR D1 D2
	TREE	0.50	69.05	92.94	32.86	68.47	91.67	30.38	A T MH PR D1 D2 D5
W4	REG	*	73.67	5.13	33.33	69.57	0.00	0.00	OZ FL D2
	LOGISTIC	0.25	63.10	54.62	62.43	50.00	89.29	63.24	FL D2
	TREE	0.25	71.36	55.46	52.52	73.91	28.57	33.33	DT OZ FL D2 D4
W5	REG	*	75.06	10.00	45.00	52.17	2.27	50.00	A FL D1 D2
	LOGISTIC	0.25	62.59	60.00	64.13	48.91	31.82	54.84	A FL D2
	TREE	0.30	73.21	61.82	52.11	55.43	18.18	38.46	A OZ FL D2

A threshold cannot be used to determine the category of symptoms in the multiple regression model.

HR, hit rate; POD, probability of detection; FAR, false alarm rate; A, autocorrelated factor; T, mean temperature; DT, daily range; MH, minimum humidity; PR, pressure; HS, hours of sunshine; OZ, concentration of ozone; PL, pollen; FL, proportion of flu patients; PM, concentration of yellow sand; YS, presence of yellow sand; D, day of the week; M, man; W, woman.

*No numerical value. **There is no significant predictive factor or dummy variable denoting day of week.

× 2 분할표를 작성하여 HR과 POD가 가능한 크고, FAR이 가능한 작게 나타나는 되는 최적 문턱치를 찾았다. (3) 성별, 연령 군별, 계절별 최종 제안 모형으로 로지스틱 회귀모형과 의사결정나무 모형 중에서 예측성 평가측도 값이 바람직한 결과를 보이는 모형을 선택하였다(Tables 5-8). (4) 검증자료에 대한 예측성 평가 측도에서 주의의 빈도가 높아진 범주 분포를 잘 따라가서 비교적 안정적이라 판단된다(Table 9). 이 범주의 결과를 성별, 연령 군별, 계절별로 제안된 예측모형들을 정리하였다.

고 찰

이번 연구에서는 환자-맞춤형 천식예보모형 개발을 목적으로, 환자의 특성을 고려하고 범주의 형태를 다양화하여 예보모형을 검증하고 개선하는 데에 중점을 두었다. 연구 과정에서 고려된 사항은 다음과 같다. (1) 환자-맞춤형을 위하여 환자의 성별과 연령을 고려하였다. (2) 천식 환자의 대체자료로서 심사평가원 약제 및 환자 자료를 사용하였다. (3) 계절에 따른 환자의 분포, 성별에 따른 환자의 분포, 연령군에 따른 환자의 분포를 조사한 결과 상이하게 나타

Table 7. Comparison of the binary forecasting models for the autumn season

Group	Model	Threshold	Skill scores (%)						Final predictors
			Whole data			Validation			
			HR	POD	FAR	HR	POD	FAR	
M1	REG	*	75.13	35.90	36.36	67.90	32.35	21.43	A T L4OZ D1
	LOGISTIC	0.30	75.61	65.98	47.11	49.38	11.76	73.33	A OZ PL D1
	TREE	0.20	80.91	81.20	36.24	59.26	64.71	48.84	A T MH PR L4OZ PM D1
M2	REG	*	80.25	43.75	25.76	77.78	34.62	10.00	A FL D1 D3-D6
	LOGISTIC	0.20	72.59	67.86	49.67	70.37	65.38	46.88	FL D1
	TREE	0.25	80.49	63.39	34.86	44.44	84.62	65.08	A MH FL D1
M3	REG	*	82.47	45.79	20.97	70.37	20.00	66.67	YS FL D1 D2
	LOGISTIC	0.20	74.07	77.57	49.39	51.85	65.00	71.11	FL D1
	TREE	0.20	82.47	75.70	35.71	75.31	0.00	0.00	A MH FL D1
M4	REG	*	70.86	8.26	41.18	67.90	0.00	100.00	PR OZ D1 D2
	LOGISTIC	0.30	65.68	52.07	56.25	69.14	60.00	50.00	PR D1 D2
	TREE	0.25	62.72	76.86	56.94	**	**	**	PR OZ D1 D2 D6
M5	REG	*	73.12	1.73	41.67	43.21	0.00	0.00	A FL PL D1 D2
	LOGISTIC	0.25	65.31	56.67	63.57	54.32	30.43	26.32	A PL D2
	TREE	0.20	79.40	73.39	39.85	46.91	26.09	42.86	A T FL PL D1 D4
W1	REG	*	79.65	41.38	21.31	71.60	36.67	26.67	A T OZ D1
	LOGISTIC	0.30	74.37	70.69	45.33	70.37	63.33	40.63	A T OZ
	TREE	0.30	83.17	71.55	29.06	65.43	40.00	45.45	A T MH PR OZ
W2	REG	*	71.85	38.17	39.76	54.32	16.28	12.50	FL D1
	LOGISTIC	0.25	54.81	81.68	59.77	45.68	65.12	50.88	A FL
	TREE	0.30	66.42	83.21	51.12	49.38	34.88	46.43	A MH FL PL D1 D4 D5
W3	REG	*	75.56	18.92	30.00	61.73	0.00	0.00	FL D1
	LOGISTIC	0.25	74.81	57.66	46.22	64.20	22.58	41.67	FL D1 D3
	TREE	0.20	77.28	50.45	39.78	64.20	22.58	41.67	A MH FL D1 D3
W4	REG	*	72.10	8.62	41.18	69.14	3.85	0.00	PR FL D1 D2
	LOGISTIC	0.25	64.69	61.21	57.99	72.84	61.54	42.86	PM FL D1 D2
	TREE	0.25	63.70	79.31	57.21	67.90	0.00	0.00	A DT PR PM FL D1 D2
W5	REG	*	73.33	2.83	62.50	46.91	0.00	0.00	FL D2 D3
	LOGISTIC	0.35	70.37	48.11	56.03	54.32	20.93	25.00	D2 D3
	TREE	0.35	76.79	52.83	44.00	51.85	13.95	25.00	A FL PL D2 D3

A threshold cannot be used to determine the category of symptoms in the multiple regression model.

HR, hit rate; POD, probability of detection; FAR, false alarm rate; A, autocorrelated factor; T, mean temperature; DT, daily range; MH, minimum humidity; PR, pressure; HS, hours of sunshine; OZ, concentration of ozone; PL, pollen; FL, proportion of flu patients; PM, concentration of yellow sand; YS, presence of yellow sand; D, day of the week; M, man; W, woman.

*No numerical value. **There is no significant predictive factor or dummy variable denoting day of week.

나 계절별, 성별, 연령군별로 각각 독립적으로 천식 범주예측모형을 개발하였다. (4) 예측 인자로는 기상 인자(기온, 습도, 일조량), 대기오염 인자(오존, 황사), 환경 인자(꽃가루), 보건 인자(독감환자비율)를 사용하였으며, 심사평가원자료의 지속성 인자와 요일 효과를 추가하였다. (5) 예보형태는 2 범주(지속관리, 주의)를 대상으로 하여 각각 예측모형을 개발한 후 결과를 비교하였다. (6) 예보대상으로 예보시간은 익일 예보를 대상으로 하였으며, 지역은 서울지역으로 하였다. (7) 예측모형으로 중회귀모형, 순서형 로지스틱모형, 범주형 의사결정나무모형을 각각 적용하고, 예측 결과를 비교하여

우수한 모형을 제안하고자 하였다. (8) 예측모형을 제안하기 위하여 전체자료를 사용한 경우의 결과와 모형검증자료(2010년도 자료)에 대한 결과를 비교하여 모형의 안정성을 살피고, 예측성 평가 척도를 사용하여 비교한 후 최적 모형을 제안하고자 하였다. 총 40 종류(4계절×2성별×5연령군)에 대하여 각각 세 모형을 적용하였으며 예측성평가에서 비교적 안정적이라 판단되어 선정된 범주예측모형을 활용하는 천식예보제를 제안한다.

천식의 치료 및 예방에서 필수적인 것은 천식 유발 인자의 노출로부터 회피하는 것이다. 천식 유발 인자에는 환자 개인의 노력으

Table 8. Comparison of the binary forecasting models for the winter season.

Group	Model	Threshold	Skill scores (%)						Final predictors
			Whole data			Validation			
			HR	POD	FAR	HR	POD	FAR	
M1	REG	*	78.59	41.23	33.80	66.00	42.11	42.86	A FL D1 D3 D4 D5 D6
	LOGISTIC	0.30	75.53	65.79	46.43	69.00	63.16	41.46	A FL D1
	TREE	0.25	82.59	60.53	29.59	71.00	39.47	28.57	A T MH FL PL D1 D2
M2	REG	*	71.29	28.57	41.54	62.00	37.50	60.00	A FL D1
	LOGISTIC	0.30	68.52	61.31	49.70	57.00	56.25	67.27	FL D1
	TREE	0.30	78.47	58.39	31.03	68.00	0.00	0.00	A T OZ FL PL D1 D5
M3	REG	*	72.69	31.16	34.85	75.00	47.22	26.09	YS FL D1 D4
	LOGISTIC	0.30	71.06	67.39	46.24	73.00	86.11	41.51	YS FL D1
	TREE	0.30	76.62	71.01	38.36	76.00	69.44	34.21	A OZ YS FL D1 D4 D5
M4	REG	*	72.92	6.56	27.27	68.00	5.88	0.00	PR OZ D2
	LOGISTIC	0.25	50.00	72.13	67.41	70.00	61.76	44.74	PR D2
	TREE	0.25	53.47	93.44	62.87	63.00	64.71	53.19	A T MH PR HS OZ D2 D3
M5	REG	*	71.53	19.05	44.19	64.00	12.50	16.67	A HS D1 D2
	LOGISTIC	0.25	60.71	79.37	58.51	71.00	70.00	37.78	A HS D1 D2
	TREE	0.25	68.71	78.57	51.71	60.00	0.00	0.00	A T PR HS OZ D1 D2
W1	REG	*	75.53	23.36	46.81	59.00	18.75	71.43	A T MH FL D1
	LOGISTIC	0.30	74.59	57.01	50.41	69.00	71.88	48.89	A T D1
	TREE	0.25	83.77	70.09	33.04	63.00	46.88	57.14	A T PR PM FL
W2	REG	*	67.82	17.48	45.65	57.00	26.67	72.41	FL D1
	LOGISTIC	0.30	67.59	61.54	49.13	51.00	70.00	65.57	FL D1
	TREE	0.30	75.93	63.64	36.36	70.00	0.00	0.00	A MH FL D1 D2 D3 D6
W3	REG	*	72.22	20.00	22.22	80.00	46.43	27.78	FL D4
	LOGISTIC	0.35	69.91	50.00	46.15	52.00	75.00	66.13	FL
	TREE	0.40	75.46	52.14	34.82	77.00	67.86	42.42	A FL D4
W4	REG	*	71.99	0.84	75.00	66.00	8.70	86.67	FL D1 D2
	LOGISTIC	0.35	66.20	40.34	60.98	59.00	21.74	82.14	D1 D2
	TREE	0.25	71.53	63.87	51.28	77.00	0.00	0.00	A FL D1 D3 D4 D5
W5	REG	*	73.41	3.54	50.00	73.00	3.70	50.00	A PR D1 D2
	LOGISTIC	0.35	67.36	42.61	60.48	71.00	22.22	57.14	D2 D4
	TREE	0.30	56.71	72.17	65.13	54.00	48.15	71.11	A MH PR D2

A threshold cannot be used to determine the category of symptoms in the multiple regression model.

HR, hit rate; POD, probability of detection; FAR, false alarm rate; A, autocorrelated factor; T, mean temperature; DT, daily range; MH, minimum humidity; PR, pressure; HS, hours of sunshine; OZ, concentration of ozone; PL, pollen; FL, proportion of flu patients; PM, concentration of yellow sand; YS, presence of yellow sand; D, day of the week; M, man; W, woman.

*No numerical value.

로 회피 가능한 인자가 있는 반면, 기상 변화, 대기오염, 꽃가루유행, 독감유행 등과 같이 개인적으로 조절 불가능한 공중보건학적 인자가 있다. 최근의 연구들에서 기상 변화나 대기오염 요인들이 천식과 유의한 연관관계가 있음을 증명되고 있다.^{17,18} 기온, 기압과 상대습도의 변화가 천식 악화에 영향이 있음을 보여주는 연구들이 있다.¹⁹⁻²¹ 천식으로 인한 입원율이 겨울에 가장 낮고 가을에 가장 높으며, 멕시코에서는 우기에 더 높다는 보고가 있다.^{21,22} 국내 한 연구에서는 기온보다 상대습도가 운동 유발성 기관지수축에 더 중요한 원인 인자임을 보고하였다.²³ 꽃가루 분포 역시 기상 요소의

영향을 받고 천식 발병에 영향을 줄 수 있다.²⁴ 대기오염 인자 특히, PM₁₀, NO₂, SO₂와 O₃의 증가는 천식발작과 연관성이 있으며, 대기오염물질이 축적되게 하는 기상 인자들의 영향이 있음이 밝혀져 있다.²⁵⁻²⁷ 바이러스 감염도 천식 악화와 연관성을 보이고,²⁸ 호흡기 바이러스의 생존 및 감염성은 강하게 기상 요소에 의존하는 양상을 보인다.²⁹

천식과 환경 요인 간의 연관성에 대한 연구 결과는 많이 있으나, 보건 예보에 중점을 둔 연구는 상대적으로 적다. 영국 기상청에서 만성 폐쇄성 호흡기 질환 환자를 대상으로 보건 예보를 시행하여

Table 9. Proposed models and thresholds for binary asthma forecasting

Group	Spring		Summer		Autumn		Winter	
	Model	Threshold	Model	Threshold	Model	Threshold	Model	Threshold
M1	LOGISTIC	0.25	LOGISTIC	0.25	TREE	0.20	LOGISTIC	0.30
M2	LOGISTIC	0.25	LOGISTIC	0.30	LOGISTIC	0.20	LOGISTIC	0.30
M3	LOGISTIC	0.25	LOGISTIC	0.40	LOGISTIC	0.20	TREE	0.30
M4	TREE	0.40	TREE	0.30	LOGISTIC	0.30	TREE	0.25
M5	LOGISTIC	0.30	LOGISTIC	0.25	TREE	0.20	LOGISTIC	0.25
W1	LOGISTIC	0.25	TREE	0.30	LOGISTIC	0.30	TREE	0.25
W2	LOGISTIC	0.25	TREE	0.30	TREE	0.30	LOGISTIC	0.30
W3	TREE	0.40	TREE	0.50	TREE	0.20	TREE	0.40
W4	TREE	0.25	TREE	0.25	LOGISTIC	0.25	LOGISTIC	0.35
W5	LOGISTIC	0.35	TREE	0.30	TREE	0.35	LOGISTIC	0.35

M, man; W, woman.

기상, 대기오염과 질환 악화와와의 연관성을 분석한 결과에 따르면, 천식예보제 시행으로 인하여 환자들이 환경 변화로 인한 질환으로 인한 위험성을 인식하고 최소화하려는 노력은 할 수 있으나 의료이 용물이나 비용 차원에서 현실적인 이득은 없었음을 알 수 있었다.^{30,31} 천식 예보와 관련해서는 영국과 그리스에서 예보모형 개발을 위한 연구 결과가 있으나 여러 가지 현실적인 제약으로 인하여 우리나라와 마찬가지로 상용화되지는 않고 있다.^{32,33}

공중보건학적 인자들의 변화와 천식 증상과의 연관성을 밝혀내고, 이를 통해 만들어진 예측모형이 천식예보제로 실용화된다면 환자 개인적인 치료에서의 이득은 물론 사회적으로 건강 증진 및 경제적 이득 효과가 있을 것이다. 환자 치료에 필요한 병원 진료 서비스의 수요를 예측할 수 있고, 치료 인력, 약제, 병원 기자재 및 설비 등이 효율적으로 환자 수요에 맞게 준비되고 적절히 분배될 것으로 예상된다.

본 연구는 심사평가원의 환자자료를 사용함으로써 대체적으로 설문지를 이용한 역학조사에 비해 의사의 진단에 따른 병명을 이용하여 신뢰할 수 있는 자료라는 점과, 서울지역 천식 환자에 대한 대표성과 객관성, 연구의 비용-효과 측면에서 장점이 있다. 그러나 천식의 진단 기준이 명확한 객관적 자료를 근거로 하지 않았다는 점, 응급실 외래로 방문한 환자만을 대상으로 하였다는 점, 검사 결과를 통한 천식 악화의 원인은 알 수 없었다는 점에서 자료 수집의 한계점을 가지고 있다. 또한, 진단적 한계점으로 현재 유병률의 급속한 증가로 인해 사회적 관심이 집중되는 알레르기 질환에 대한 의사의 과진단을 배제할 수 없다. 본 연구는 서울지역에 국한된, 6년간의 자료를 분석한 것이었기에 이를 통해 도출된 예보모형을 보편화하기에는 무리가 있다. 반면, 본 연구에서는 외국의 연구 결과에서는 없었던, 호흡기 질환의 악화에 영향을 주는 흔하고, 중요한 인자인 인플루엔자 바이러스의 유행과 꽃가루 자료가 예측 인자로써 함께 포함되어 있다.

결론적으로, 본 연구에서는 천식예보모형으로 이 범주 예측모형의 적합성을 제시하였다. 최근 기후 변화, 알레르기 질환의 증가 현상 등으로 인하여 알레르기 환자들뿐 아니라 일반 국민들에게도 우리나라 환경과 천식 질환자의 특성에 맞는 환자 맞춤형 천식예보제의 개발 및 활용은 환자 개인의 건강 증진 및 공공 의료의 질 향상을 위해 절실하게 필요하다. 이를 위하여 환자 및 예측 인자들에 대한 지속적인 관찰과 예보제 보완 및 상용화를 위한 정부의 협조적인 정책이 필요할 것으로 생각한다.

REFERENCES

- Fanta CH. Asthma. *N Engl J Med* 2009;360:1002-14.
- Braman SS. The global burden of asthma. *Chest* 2006;130(1 Suppl):4S-12S.
- Kroegel C. Global Initiative for Asthma (GINA) guidelines: 15 years of application. *Expert Rev Clin Immunol* 2009;5:239-49.
- Masoli M, Fabian D, Holt S, Beasley R; Global Initiative for Asthma (GINA) Program. The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy* 2004;59:469-78.
- Choi JH, Cha YM. The prevalence of allergic diseases among Korean children. *Korean J Pediatr* 1964;7:39-41.
- Lee HR, Hong DS, Sohn KC. Survey of allergic diseases in children. *J Korean Med Assoc* 1983;26:254-62.
- Shin TS, Lee GJ, Yoon HS. A survey of the distribution of allergic diseases in primary school children. *Korean J Asthma Allergy Clin Immunol* 1990;10:201-12.
- Kim YK, Kim SH, Tak YJ, Jee YK, Lee BJ, Kim SH, et al. High prevalence of current asthma and active smoking effect among the elderly. *Clin Exp Allergy* 2002;32:1706-12.
- The Global Initiative for Asthma. The Global Asthma Report 2014. Global burden of disease due to asthma [Internet]. The Global Initiative for Asthma; [2015 Sep 15]. Available from: <http://www.globalasthmareport.org/burden/burden.php>.
- Beasley R. The burden of asthma with specific reference to the United States. *J Allergy Clin Immunol* 2002;109(5 Suppl):S482-9.

11. Kim CY, Park HW, Ko SK, Chang SI, Moon HB, Kim YY, et al. The financial burden of asthma: a nationwide comprehensive survey conducted in the republic of Korea. *Allergy Asthma Immunol Res* 2011;3:34-8.
12. A fresh perspective on asthma. *Nat Med* 2012;18:631.
13. Pedersen SE, Hurd SS, Lemanske RF Jr, Becker A, Zar HJ, Sly PD, et al. Global strategy for the diagnosis and management of asthma in children 5 years and younger. *Pediatr Pulmonol* 2011;46:1-17.
14. O'Connor GT. Allergen avoidance in asthma: what do we do now? *J Allergy Clin Immunol* 2005;116:26-30.
15. Soyiri IN, Reidpath DD. Evolving forecasting classifications and applications in health forecasting. *Int J Gen Med* 2012;5:381-9.
16. Tobías A, Sáez M, Galán I, Campbell MJ. Sensitivity analysis of common statistical models used to study the short-term effects of air pollution on health. *Int J Biometeorol* 2003;47:227-9.
17. Ivey MA, Simeon DT, Monteil MA. Climatic variables are associated with seasonal acute asthma admissions to accident and emergency room facilities in Trinidad, West Indies. *Clin Exp Allergy* 2003;33:1526-30.
18. Chavarría JF. Short report: Asthma admissions and weather conditions in Costa Rica. *Arch Dis Child* 2001;84:514-5.
19. Carey MJ, Cordon I. Asthma and climatic conditions: experience from Bermuda, an isolated island community. *Br Med J (Clin Res Ed)* 1986;293:843-4.
20. Khot A, Burn R, Evans N, Lenney W, Storr J. Biometeorological triggers in childhood asthma. *Clin Allergy* 1988;18:351-8.
21. Garty BZ, Kosman E, Ganor E, Berger V, Garty L, Wietzen T, et al. Emergency room visits of asthmatic children, relation to air pollution, weather, and airborne allergens. *Ann Allergy Asthma Immunol* 1998;81:563-70.
22. Rosas I, McCartney HA, Payne RW, Calderón C, Lacey J, Chapala R, et al. Analysis of the relationships between environmental factors (aeroallergens, air pollution, and weather) and asthma emergency admissions to a hospital in Mexico City. *Allergy* 1998;53:394-401.
23. Koh YI, Choi IS. Seasonal difference in the occurrence of exercise-induced bronchospasm in asthmatics: dependence on humidity. *Respiration* 2002;69:38-45.
24. Gioulekas D, Balafoutis C, Damialis A, Papakosta D, Gioulekas G, Patakas D. Fifteen years' record of airborne allergenic pollen and meteorological parameters in Thessaloniki, Greece. *Int J Biometeorol* 2004;48:128-36.
25. Jalaludin BB, O'Toole BI, Leeder SR. Acute effects of urban ambient air pollution on respiratory symptoms, asthma medication use, and doctor visits for asthma in a cohort of Australian children. *Environ Res* 2004;95:32-42.
26. Holmén A, Blomqvist J, Frindberg H, Johnelius Y, Eriksson NE, Henriksen KA, et al. Frequency of patients with acute asthma in relation to ozone, nitrogen dioxide, other pollutants of ambient air and meteorological observations. *Int Arch Occup Environ Health* 1997;69:317-22.
27. Anderson HR, Ponce de Leon A, Bland JM, Bower JS, Emberlin J, Strachan DP. Air pollution, pollens, and daily admissions for asthma in London 1987-92. *Thorax* 1998;53:842-8.
28. Johnston SL, Pattemore PK, Sanderson G, Smith S, Campbell MJ, Josephs LK, et al. The relationship between upper respiratory infections and hospital admissions for asthma: a time-trend analysis. *Am J Respir Crit Care Med* 1996;154(3 Pt 1):654-60.
29. Hersoug LG. Viruses as the causative agent related to 'dampness' and the missing link between allergen exposure and onset of allergic disease. *Indoor Air* 2005;15:363-6.
30. Maheswaran R, Pearson T, Hoysal N, Campbell MJ. Evaluation of the impact of a health forecast alert service on admissions for chronic obstructive pulmonary disease in Bradford and Airedale. *J Public Health (Oxf)* 2010;32:97-102.
31. Bakerly ND, Roberts JA, Thomson AR, Dyer M. The effect of COPD health forecasting on hospitalisation and health care utilisation in patients with mild-to-moderate COPD. *Chron Respir Dis* 2011;8:5-9.
32. Soyiri IN, Reidpath DD, Sarran C. Forecasting peak asthma admissions in London: an application of quantile regression models. *Int J Biometeorol* 2013;57:569-78.
33. Moustris KP, Douros K, Nastos PT, Larissi IK, Anthracopoulos MB, Paliatsos AG, et al. Seven-days-ahead forecasting of childhood asthma admissions using artificial neural networks in Athens, Greece. *Int J Environ Health Res* 2012;22:93-104.