Review Article

J Korean Soc Radiol 2024;85(4):714-726
https://doi.org/10.3348/jksr.2024.0050
eISSN 2951-0805

JOURNAL of THE KOREAN SOCIETY of RADIOLOGY

Using Artificial Intelligence Software for Diagnosing Emphysema and Interstitial Lung Disease

폐기종 및 간질성 폐질환: 인공지능 소프트웨어 사용 경험

Sang Hyun Paik, MD[1], Gong Yong Jin, MD[2]*

[1]Department of Chest, Korea Teleradiology Reading Center, Seoul, Korea
[2]Department of Radiology, Research Institute of Clinical Medicine of Jeonbuk National University-Biomedical Research Institute of Jeonbuk National University Hospital, Jeonbuk National University and Medical School, Jeonju, Korea

Received April 11, 2024
Revised June 23, 2024
Accepted July 18, 2024

***Corresponding author**
Gong Yong Jin, MD
Department of Radiology, Research Institute of Clinical Medicine of Jeonbuk National University-Biomedical Research Institute of Jeonbuk National University Hospital, Jeonbuk National University and Medical School,
20 Geonji-ro, Deokjin-gu, Jeonju 54907, Korea.

Tel 82-63-250-2307
E-mail gyjin@chonbuk.ac.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Researchers have developed various algorithms utilizing artificial intelligence (AI) to automatically and objectively diagnose patterns and extent of pulmonary emphysema or interstitial lung diseases on chest CT scans. Studies show that AI-based quantification of emphysema on chest CT scans reveals a connection between an increase in the relative percentage of emphysema and a decline in lung function. Notably, quantifying centrilobular emphysema has proven helpful in predicting clinical symptoms or mortality rates of chronic obstructive pulmonary disease. In the context of interstitial lung diseases, AI can classify the usual interstitial pneumonia pattern on CT scans into categories like normal, ground-glass opacity, reticular opacity, honeycombing, emphysema, and consolidation. This classification accuracy is comparable to chest radiologists (70%–80%). However, the results generated by AI are influenced by factors such as scan parameters, reconstruction algorithms, radiation doses, and the training data used to develop the AI. These limitations currently restrict the widespread adoption of AI for quantifying pulmonary emphysema and interstitial lung diseases in daily clinical practice. This paper will showcase the authors' experience using AI for diagnosing and quantifying emphysema and interstitial lung diseases through case studies. We will primarily focus on the advantages and limitations of AI for these two diseases.

Index terms Artificial Intelligence; Emphysema; Interstitial Lung Disease; Computed Tomography

서론

흉부의 전산화단층촬영(이하 CT) 또는 흉부 X선을 이용한 의료 영상에서 인공지능

Copyrights © 2024 The Korean Society of Radiology

(artificial intelligence; 이하 AI)은 현재 다양한 분야에서 실험적 또는 임상적으로 활용되고 있다. 의료 영상 분야에서 인공지능의 발달은 기존의 머신 러닝보다 성능이 우수한 딥 러닝(deep learning) 방법, 특히 심층 합성공 신경망(convolutional neural networks; 이하 CNN)의 개발 때문이다. 이와 더불어 인공지능을 학습시킬 수 있는 많은 데이터가 공급이 되면서부터 이를 활용하여 흉부 CT나 흉부 X선을 이용한 다양한 진단(예를 들면, 폐결절의 발견과 폐암 가능성 진단, 폐렴과 기흉 진단, 폐기종의 정량화, 간질성 폐질환의 정량화)에 활용이 가능해졌다. 이 질환들 중에 흉부 CT에서 폐기종과 간질성 폐질환(interstitial lung disease; 이하 ILD)은 단순히 진단을 하는 것 외에도 병변의 형태나 범위가 환자의 폐 기능 평가나 예후에 밀접한 관계가 있어서 병변의 정량적인 분석을 시도해 왔다(1, 2).

흉부 CT에서 주관적으로 폐기종과 ILD를 진단하고, 형태(pattern)를 분류하고, 범위(extent)를 알 수는 있지만 이 질환들을 객관적으로 형태를 분류하고 범위를 결정하는 데 한계가 있다. 지금까지 흉부 CT상 폐기종이나 ILD의 형태나 범위의 평가는 시각적인 방법(visual assessment)과 정량화(quantification) 분석방법들이 광범위하게 시도되어 왔고, 그 유용성도 증명이 되었지만 여전히 이런 방법으로는 분석을 하는 데 시간이 많이 걸리고, 주관적이며, 일정한 결과를 얻기 위해서는 연구자들을 지속적으로 훈련을 시켜야 하기 때문에 일상적인 업무에서는 거의 활용할 수 없는 단점이 있다. 이런 문제점을 해결하고자 흉부 CT에서 발견된 폐기종이나 간질성 폐질환의 형태나 범위를 인공지능을 이용하여 자동적으로 객관적으로 분류하는 다양한 알고리즘을 개발하고, 이를 증명하는 연구들이 진행되고 있다(1-5).

그러나 아직까지 흉부 CT상 폐기종과 ILD의 패턴 분석과 병변의 범위 정량화를 위해 인공지능을 실제로 일상 업무에서 흔히 사용하고 있지는 않다. 이 논문에서는 폐기종과 ILD의 패턴 분석과 정량화를 위해서 AI를 사용하고 있는 저자들의 경험을 증례로 소개를 하고, 이 두 질환의 AI 효용성과 제한점에 대해서 언급하고자 한다.

폐기종과 간질성 폐질환에 대한 AI 소프트웨어

폐기종 정량화 AI

폐기종은 폐포벽(alveolar wall)의 비가역적 파괴로 인해 원위부 공기공간(distal airspace)이 확장되는 질환이다. 비록 폐기종이 처음에는 폐 섬유화를 일으키지는 않지만 말기 폐기종으로 진행하면 결국 비가역적인 폐 섬유화를 유도하기도 한다. 폐기종을 포함한 만성 폐쇄성 폐질환은 2009년에서 2019년 사이에 사망률이 35.4% 증가되었으며, 2030년까지 세계 사망률의 3번째 원인이 될 것으로 예측하고 있다(6). 이로 인해서 여전히 공중 보건(public health)에 큰 부담이 되고 있다. 따라서, 이 질환을 효과적으로 진단하고 적절한 치료 전략을 수립하는 것은 이 질환으로 인한 사망률을 감소시키기 위해 매우 중요하다(7). 폐기종의 가장 중요한 두 가지 원인은 흡연(smoking)과 α1-antitrypsin-deficiency이다. 우리나라의 경우에는 흡연이 가장 중요한 원인이며, 흡연으로 인한 폐기종의 경우 주로 양측 상엽에 발생한다. 폐기종을 포함한 만성폐쇄성 폐질환을 치료하고 지속적인 관리를 위해서 폐기종의 분포와 정량화가 중요하기 때문에 흉부 CT로 폐

기종을 진단하고, 이를 정량화하는 것은 매우 중요하다(7, 8).

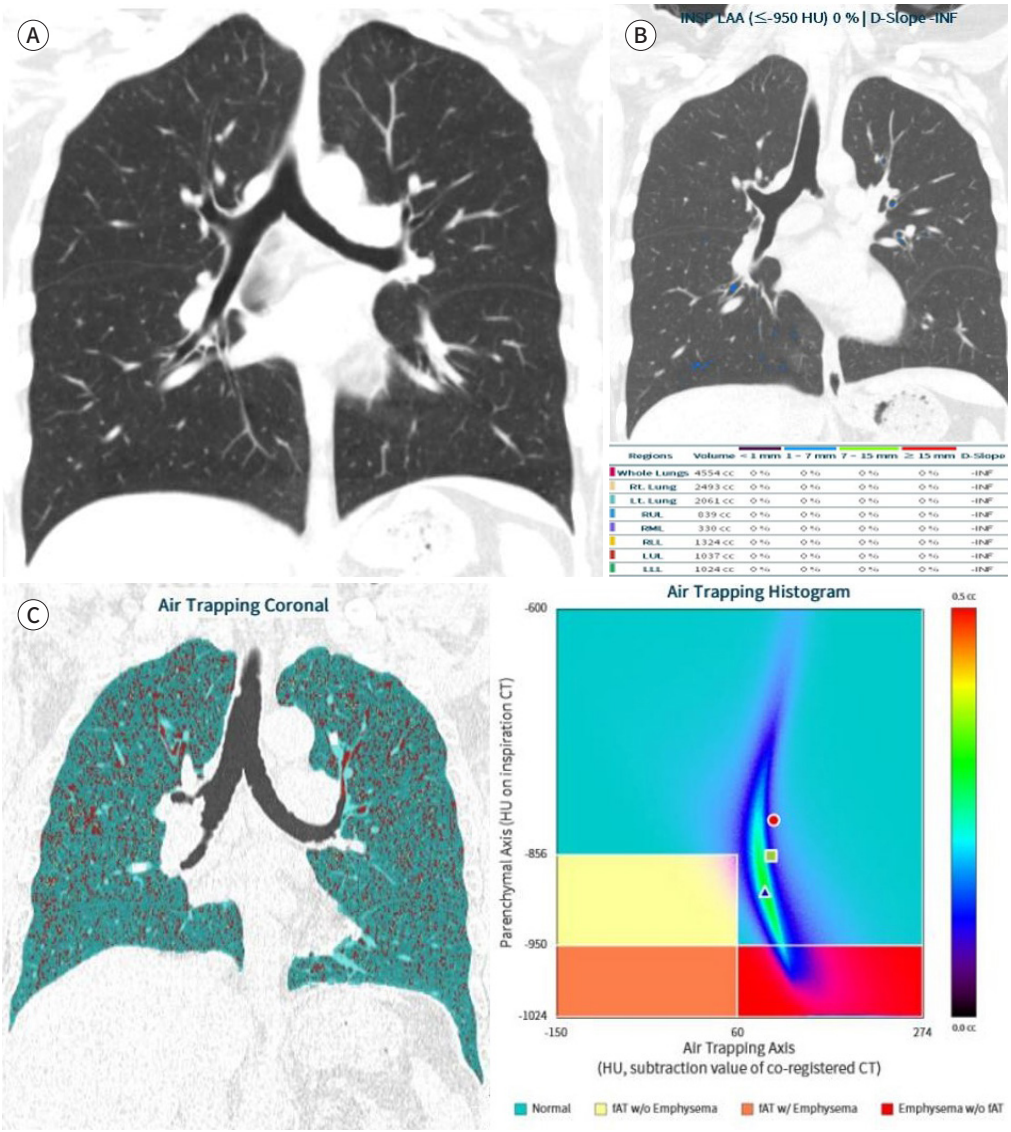
흉부 CT는 폐기종이 있는 폐를 쉽게 진단할 수 있고, 폐기종의 분포와 정량화를 할 수 있다. 흉부 CT에서 폐기종을 정량화하는 방법으로는 시각적 채점(visual scoring), 농도계(densitometry)를 사용한 반정량적 방법(semiquantitative method) 또는 정량화(quantification) 방법이 있다. 비록 흉부 CT에서 폐기종의 범위를 시각적 채점으로 하는 방법이 폐기종의 진행을 평가하는 데 유용한 방법으로 알려져 있지만, 이 방법은 측정 방법의 표준화가 어렵고, 폐기종을 평가하는 데 시간이 많이 걸리며, 평가자들이 일정하게 폐기종을 평가하도록 교육하는 것이 어렵고, 시간이 많이 걸린다. 이와 더불어 평가자들 사이의 불일치도 33% 정도로 매우 높다(1). 이런 이유로 인공신경망을 이용한 딥 러닝 인공지능 기술을 이용하여 폐기종을 자동으로 정량화할 수 있는 상업적인 소프트웨어가 폐기종을 정량화하는 방법이 좋은 진단법(Global Initiative for Chronic Obstructive Lung Disease [이하 GOLD] standard)으로 받아들여지고 있다(1, 2, 7, 9).

폐기종은 흉부 CT에서 정상 감쇄(attenuation)를 가진 폐 실질로 둘러싸여 있는 저 감쇄 영역(low attenuation area)으로 보인다. 이런 차이를 이용하여 폐기종을 정량화하는데, 현재 상용되고 있는 폐기종 정량화 소프트웨어의 기술적인 방법은 동일하다. 첫 번째 단계는 폐 실질을 흉벽 및 종격동과 같은 다른 해부학적 구조들을 가상적으로 제거한다. 두 번째 단계는 남아 있는 폐 실질을 복셀(voxel)단위로 밀도를 측정하는 데 기준은 하운스필드 유닛(Hounsfield unit; 이하 HU)이다. 흉부 CT에서 정상 폐 실질과 폐기종을 구분할 수 있는 HU의 임계값은 -950 HU로 받아들여지고 있으며(7), 이를 기준으로 정상 폐와 폐기종을 구분한 후 자동적으로 폐기종과 정상 폐의 용량을 측정한다. 마지막으로 -950 HU보다 낮게 측정되는 복셀을 자동적으로 추출하여 정상 폐 실질과 비교해 폐기종의 상대적인 양(relative percentage)을 표시한다(7-9).

이러한 기법을 이용하여 흉부 CT상 폐기종을 정량화하고, 폐 기능 검사와 비교하여 임상적인 유용성에 대한 연구들이 있지만, 여전히 임상적인 적용에는 제한점이 있다(2, 3, 7). 폐기종 정량화 연구들을 보면 흉부 CT상 폐기종을 다양한 프로그램을 이용해서 정량화를 하고 있는 데, 공통적인 결과로는 폐기종의 상대적인 양의 증가가 폐 기능의 악화와 연관이 있다고 언급한다(7, 9, 10). 예를 들면 Sileikienė 등(10)은 GOLD stage 1 and 2 (forced expiratory volume in one second [이하 FEV1] = mean 63%) 그룹과 GOLD stage 3 and 4 (FEV1 = mean 39%) 그룹에서 폐기종으로 생각되는 -950 HU (%)의 영역을 비교해 보면 각각 6%와 15%로 통계적으로 의미 있게 차이가 있다고 했다. 저자의 경우에도 상용화된 인공지능 소프트웨어를 이용하여 흉부 CT에서 폐기종을 정량화했을 때 정상인 경우와 폐기종이 심한 경우 -950 HU (%)에 차이가 나는 것을 경험하였고, 임상의에게 폐기종 환자의 흉부 CT를 정량화한 후 -950 HU (%)의 수치를 제시했을 때 환자의 증상이나 폐 기능과 연관성이 있음을 알 수 있었다(Figs. 1, 2). 그러나 여전히 폐기종의 정량화를 임상적으로 적용하기에는 몇 가지 제한점이 있다. 흉부 CT에서 폐기종의 정량화는 스캔 변수들(scan parameters), 재구성 알고리즘(reconstruction algorithm), 그리고 방사선 선량(radiation dose)에 영향을 받기 때문에 CT상 이런 조건들이 동일하지 않으면 -950 HU (%)의 수치를 그대로 모든 병원에서 임상에 적용하는 것은 바람직하지 않다고 생각한다(7, 10-12) (Figs. 3-5). 비록 상용화된 인공지능 폐기종 정량화 프로그램마다 요구하는 흉부 CT 촬영 조건이 있고, 이런 촬영 조건

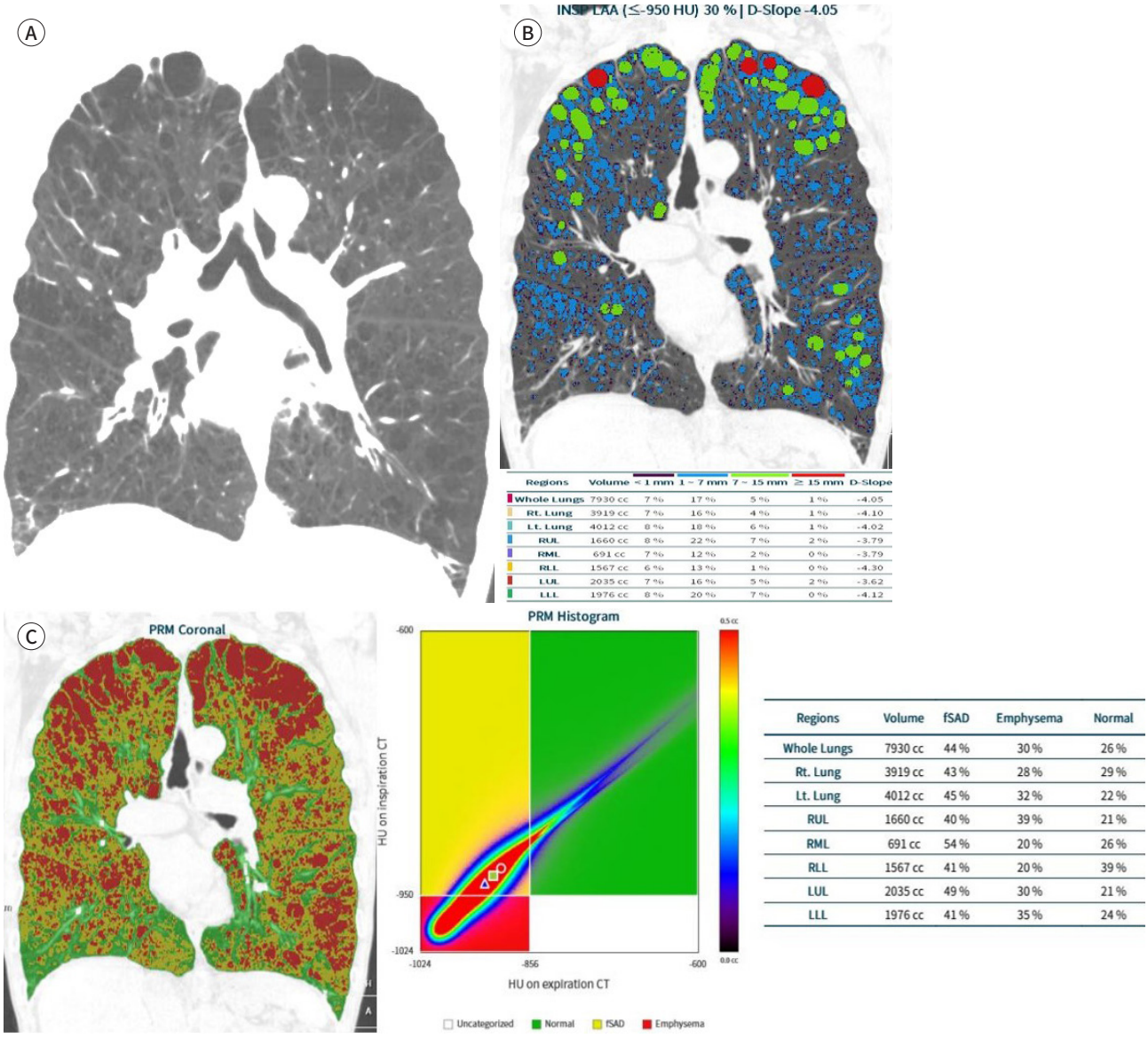
JOURNAL of
THE KOREAN SOCIETY of
RADIOLOGY

Fig. 1. A 39-year-old male patient never smoker underwent chest CT. Pulmonary function tests showed normal FEV1/FVC.
A. The inspiratory chest CT scan shows no emphysema.
B. Quantitative analysis reveals a 0% low attenuation area.
C. Air trapping mapping using paired inspiratory and co-registered chest CT scans demonstrates 0% air trapping in both lungs.
FEV1 = forced expiratory volume in one second, FVC = forced vital capacity, HU = Hounsfield unit



으로 얻어진 흉부 CT를 인공지능을 이용해서 분석을 했을 때 정확한 폐기종 정량화 결과를 얻을 수 있지만, 각 병원별로 다양한 변수를 고려해서 일정하게 흉부 CT를 촬영한 후 폐기종 정량화 결과를 임상에 적용하는 것이 바람직하다. 또한, 위에서 언급했던 -950 HU (%)를 기준으로 폐기종을 정량화하는 방법이 일부 환자에서 정량화 수치와 환자의 임상 증상과 잘 맞지 않기도 하다. 다양한 이유가 있겠지만 지금까지 개발된 폐기종 정량화 인공지능 소프트웨어가 폐기종의 형태를 구분하지 못하는 것도 그 원인 중에 하나이다. 폐기종의 형태 중 중심성 폐기종(centrilobular emphysema)과 범소엽성 폐기종(panlobular emphysema)은 환자의 호흡기 증상과 밀접한 관

Fig. 2. A 74-year-old male patient current smoker (50 pack-years) with severe shortness of breath underwent chest CT. Pulmonary function tests showed a reduced FEV1/FVC of 68%.
A. Chest CT reveals severe empyema throughout the lungs, with centrilobular and panlobular emphysema with bullae predominantly distributed in both upper lobes.
B. Quantitative analysis of the inspiratory chest CT scan shows a 30% low attenuation area (colored). Notably, empyema with a size of 1–7 mm occupies the majority of this area.
C. Parametric response mapping using paired inspiratory and expiratory chest CT scans demonstrates 44% SAD (fSAD) in both lungs.
FEV1 = forced expiratory volume in one second, fSAD = functional small airway disease, FVC = forced vital capacity, HU = Hounsfield unit



Regions	Volume	< 1 mm	1 ~ 7 mm	7 ~ 15 mm	≥ 15 mm	D-Slope
Whole Lungs	7930 cc	7 %	17 %	5 %	1 %	-4.05
Rt. Lung	3919 cc	7 %	16 %	4 %	1 %	-4.10
Lt. Lung	4012 cc	8 %	18 %	6 %	1 %	-4.02
RUL	1660 cc	8 %	22 %	7 %	2 %	-3.79
RML	691 cc	7 %	12 %	2 %	0 %	-3.79
RLL	1567 cc	6 %	13 %	1 %	0 %	-4.30
LUL	2035 cc	7 %	16 %	5 %	2 %	-3.62
LLL	1976 cc	8 %	20 %	7 %	0 %	-4.12

Regions	Volume	fSAD	Emphysema	Normal
Whole Lungs	7930 cc	44 %	30 %	26 %
Rt. Lung	3919 cc	43 %	28 %	29 %
Lt. Lung	4012 cc	45 %	32 %	22 %
RUL	1660 cc	40 %	39 %	21 %
RML	691 cc	54 %	20 %	26 %
RLL	1567 cc	41 %	20 %	39 %
LUL	2035 cc	49 %	30 %	21 %
LLL	1976 cc	41 %	35 %	24 %

계가 있지만 중격주위 폐기종(paraseptal emphysema)은 임상적으로 거의 중요하지 않다(7, 13). 따라서 현재 개발된 폐기종 정량화 상용소프트웨어가 폐기종의 형태를 구분하지 못하기 때문에 이런 소프트웨어에서 제시한 폐기종의 정량화 수치를 모든 환자에서 임상적으로 의미를 갖기에는 제한점이 있다.

폐기종의 원인은 흡연이 가장 흔하기 때문에 흡연으로 인한 폐기종, 즉 중심성 폐기종의 중증도

Fig. 3. A 61-year-old male patient current smoker (30 pack-years) underwent low-dose chest CT.
A. The low-dose chest CT scan shows no emphysema.
B. Quantitative analysis of the inspiratory low-dose chest CT scan reveals a 0% low attenuation area.

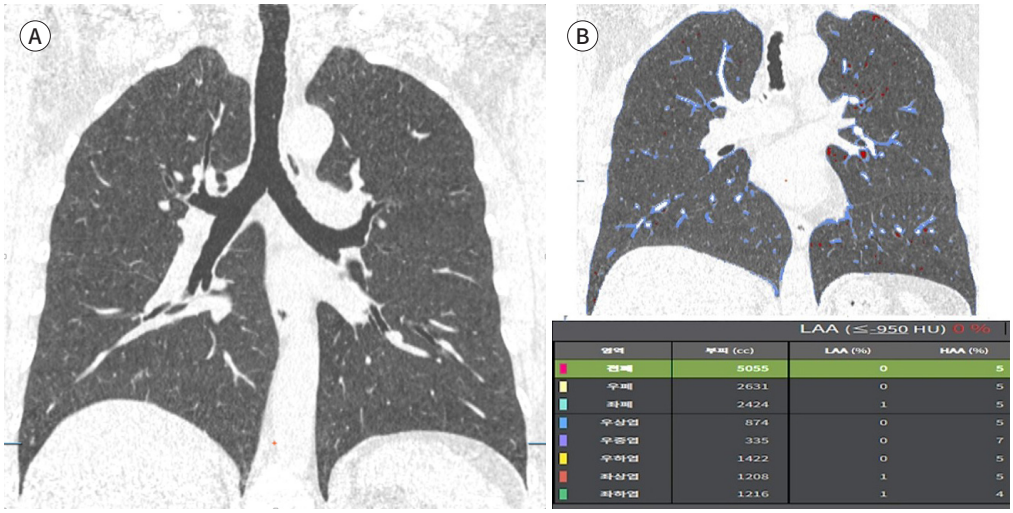
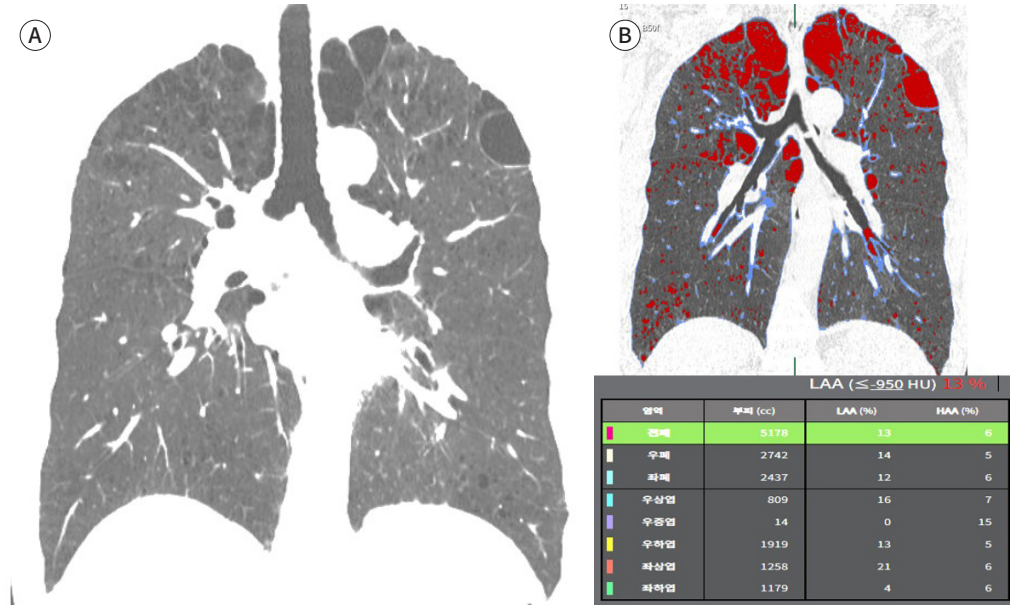
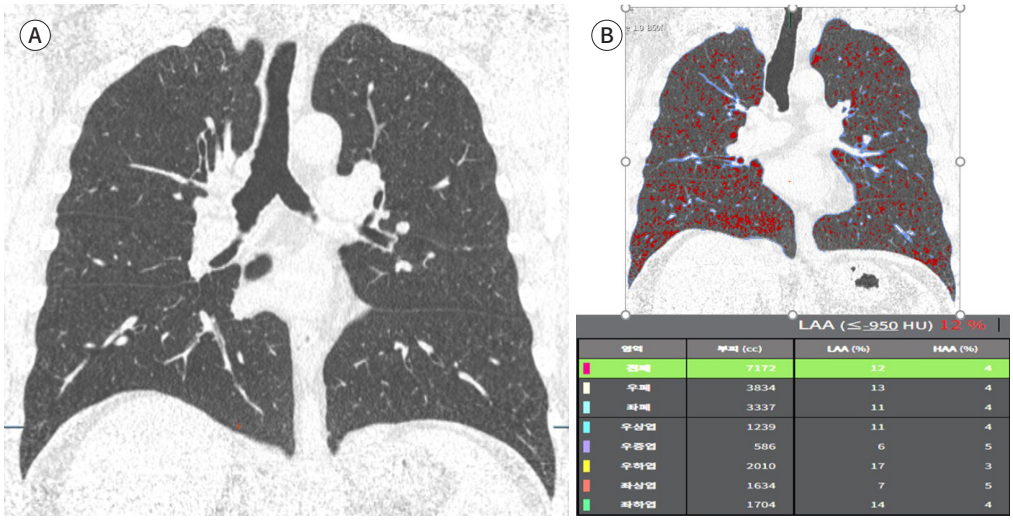


Fig. 4. A 67-year-old male patient current smoker (30 pack-years) underwent low-dose chest CT.
A. The inspiratory low-dose chest CT scan shows centrilobular empyema throughout the lungs, with emphysema with bullae predominantly distributed in both upper lobes.
B. Quantitative analysis of the low-dose chest CT scan shows a 13% low attenuation area (colored in red).



를 흉부 CT에서 알려주는 것이 임상적으로 도움이 된다. 그러나 위에서 언급했던 것처럼 현재까지 상용화된 폐기종 정량화 소프트웨어는 폐기종의 형태를 구분할 수 없기 때문에 결국 시각적으로 중심성 폐기종을 중심으로 정량화를 하는 것이 임상 증상이나 만성폐쇄성 폐질환의 사망률을 예측하는 데 도움이 된다(1, 8, 9, 11, 14). Fleischner Society는 폐기종의 중증도를 반영할 수 있는 CT 근거 시각 분류(CT based visual classification)를 제안했는데, 중심성 폐기종을 심한 정도에 따라서 흔적(trace), 경미(mild), 중등도(moderate), 융합(confluent), 진행성 파괴(advanced de-

Fig. 5. A 61-year-old male patient ex-smoker underwent low-dose chest CT.
A. The low-dose inspiratory chest CT scan shows no emphysema.
B. Quantitative analysis of the low-dose chest CT scan reveals a 12% low attenuation area (colored in red). This is an artifact due to the low-dose chest CT.



struction)로 구분하였다(14). El Kaddouri 등(1)이 COPDGene 참가자(n = 4166) 중 만성폐쇄성폐질환으로 진단된 1665명(40%) 참가자를 대상으로 폐기종을 CT 근거 시각 분류를 이용하여 폐기종 중증도와 폐기종의 진행의 상관관계에 대해서 연구했을 때 처음 CT에서 폐기종이 있는 경우 5년 후 추적 CT에서 폐기종은 진행이 되었지만 처음 CT에서 폐기종이 없는 경우에 폐기종의 진행은 없었으며, 폐기종의 중증도가 심할수록 폐기종의 진행 속도가 더 빨랐다고 했다. 그러나 이런 방법도 시각적 채점 방법이 가지고 있던 한계를 극복할 수는 없다.

최근에 CT상 폐기종을 정량화하기 위한 방법들 중 -950 HU (%)를 이용한 정량화와 기존의 시각적 채점 방법의 단점을 보완하는 AI 기법이 제시가 되었다. 폐기종의 정량화를 -950 HU (%)가 아닌 Fleischner Society가 제시한 CT 근거 폐기종 분류를 이용한 AI 폐기종 정량화 방법이다. Humphries 등(3)은 COPDGene 참가자의 CT와 Fleischner Society에서 제시한 CT 근거 폐기종 분류를 이용해서 폐기종의 중증도를 분류할 수 있는 자동 AI 알고리즘을 개발한 후 COPDGene 참가자를 대상으로 유용성에 대해서 검증을 하였는데, 개발된 자동 AI 알고리즘이 시각 분류 방법보다 더 객관적으로 평가를 할 수 있었으며, 흔적(trace) 형태는 시각 분류 방법보다 더 잘 발견할 수 있다고 했다. 이 방법이 기존의 시각 분류 방법이 가지고 있던 단점을 개선할 수 있다고 주장했다. 그러나 이 방법도 흉부 CT의 스캔 변수들(scan parameters), 재구성 알고리즘(reconstruction algorithm), 그리고 방사선 선량(radiation dose)이 영향을 주는지와 COPDGene 참가자 외의 데이터로 검증하는 추가적인 연구가 필요하다.

간질성 폐질환 정량화 AI

특발성 간질성 폐질환(idiopathic interstitial pneumonia; 이하 IIP)에는 만성 섬유화성 간질성 폐렴(특발성 폐 섬유증[idiopathic pulmonary fibrosis; 이하 IPF])와 특발성 비특이적 간질성 폐

렴(nonspecific interstitial pneumonia), 흡연 관련 간질성 폐렴(smoking-related interstitial pneumonia), 급성 또는 아급성 간질성 폐렴(acute or subacute interstitial pneumonia), 그리고 급성 간질성 폐렴(acute interstitial pneumonia)이 있다. IIP의 가장 흔한 원인은 만성 진행성 섬유성 간질성 폐질환(chronic progressive fibrotic ILD)이며, 통상성 간질성 폐렴(usual interstitial pneumonia; 이하 UIP)이란 용어로 사용되어 왔다(15). 이 질환은 흉부 CT에서 양측 하엽의 늑막하(subpleural area)에 대칭적으로 위치하며, 특징적인 CT 소견으로는 벌집 모양(honeycomb), 망상형 음영(reticular opacity), 견인성 기관지와 세기관지 확장증(traction bronchiectasis and bronchiolectasis)을 보인다. CT상 명확한(definitive) UIP 형태는 병리학적 UIP에 대해서 90% 이상의 높은 양성예측도(positive predictive value)를 보이기 때문에 수술적 폐 조직검사를 하지 않고 흉부 CT로만 진단이 가능하다(15).

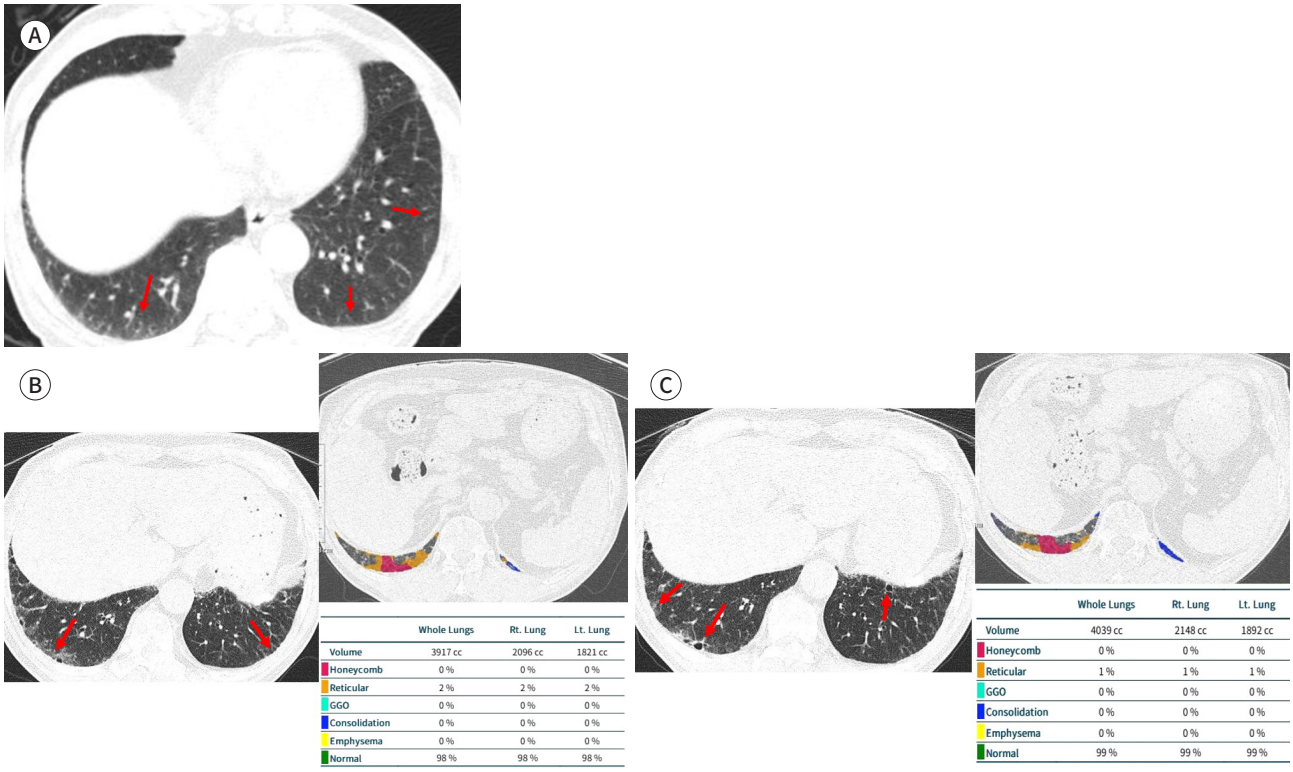
비록 흉부 CT에서 IPF를 진단하는 중요한 소견이 벌집 모양과 망상형 음영이지만 이러한 소견이 IPF와 폐 섬유화를 일으키는 다른 질환을 CT만으로 감별을 할 수는 없다. 그 이유로는 IPF에서 보이는 UIP와 다른 원인(결합조직병[connective tissue disease], 만성과민성폐렴[chronic hypersensitivity pneumonia], 진폐증[pneumoconiosis] 등)에서 보이는 UIP를 CT로 감별하는 것이 불가능하기 때문이다(16). 그러나 원인이 무엇이든지 처음 흉부 CT에서 망상형 음영과 주변에 간유리음영(ground glass opacity)이 발견이 되고, 이러한 소견이 폐 섬유화로 진행이 된다면 망상형 음영의 범위 증가, 폐 실질의 구조적 비틀림(architecture distortion), 견인성 기관지와 세기관지 확장증, 그리고 결국 벌집 모양의 폐 섬유화 순서로 진행이 되므로 환자의 치료나 예후에 IPF의 흉부 CT 모니터링은 매우 중요하다(15). 흉부 CT는 IPF 환자에게 조직 검사를 할 것인지 또는 항염증제(anti-inflammatory drug)나 항섬유화제(antifibrotic drug)를 투여할 것인지를 결정하는 데 매우 중요한 역할을 하며, 이와 더불어 망상형 음영과 벌집 모양의 범위를 반영하는 섬유화 점수(fibrotic score)나 견인성 기관지 확장증과 세기관지 확장증 지표(traction bronchiectasis/bronchiolectasis index)를 이용한 정량화가 환자의 예후를 예측하는 데 도움이 될 수 있다(15, 17-19).

흉부 CT에서 UIP 형태를 시각적 채점(visual scoring)을 이용해서 정량화를 할 경우 측정 방법의 표준화가 어렵고, 평가하는 데 시간이 많이 걸리며, 평가자들이 UIP의 형태를 분류하는 데 시간이 많이 걸리고, 평가자들 간의 일치도 또한 높지 않다. 이와 같은 시각적 채점의 단점을 극복하고자 AI를 이용하여 CT에서 UIP 형태를 정량화하는 시도들이 있다(4, 20). 폐기종의 AI 정량화에 과정과 비교했을 때 UIP 형태의 정량화는 더 복잡한 분석 단계를 거친다. 먼저 폐 실질을 다양한 해부학적 구조물로부터 분할을 한 후 폐 실질로부터 UIP가 있는 폐 실질만을 검출한다. 다음 단계로 딥 특징(deep features) 또는 레디오믹 특징(radiomic features)를 이용하여 UIP 형태, 즉 벌집 모양, 망상형 음영 등을 추출하고, 정량화를 한다. 마지막으로 데이터를 다시 분석해서 UIP 형태를 전체 폐 실질에서 차지하는 양을 정량화하거나 변화를 임상적으로 사용할 수 있도록 제공한다(Figs. 6-8) (20).

지금까지 흉부 CT에서 UIP 형태 영상 분석 AI 모델들은 주로 UIP의 형태를 분할(segmentation) 또는 분류(classification)하는 모델들이었으며, 최근에는 진행성 섬유성 간질성 폐질환(progressive fibrotic ILD) 형태를 정량화(quantification)까지 하는 AI 모델들도 개발되고 있다(2-5,

JOURNAL of THE KOREAN SOCIETY of RADIOLOGY

Fig. 6. A 74-year-old male patient current smoker (30 pack-years) underwent low-dose chest CT.
A. The patient underwent the low-dose chest CT for initial lung cancer screening at a time when he had no respiratory symptoms. The initial low-dose CT scan shows multifocal traction bronchiolectasis (arrows) in the subpleural area of both lungs.
B. Four years later, the patient developed dyspnea and underwent another chest CT scan. A lung biopsy confirmed the usual interstitial lung pneumonia. The chest CT scan shows diffuse GGA with traction bronchiolectasis and interlobular septal thickening (arrows) in the subpleural area of both lungs. Compared to a low-dose chest CT scan from four years ago, there is a progression of subpleural fibrosis in both lungs. Quantitative analysis of the chest CT scan shows 2% reticular opacity (orange color) and less than 1%, honeycomb (red color).
C. The patient was treated with Pirespa tablets (200 mg) for one year. A follow-up chest CT scan shows GGA with traction bronchiolectasis and interlobular septal thickening (arrows) in the subpleural area of both lungs. Compared to the chest CT from one year ago, subpleural fibrosis in both lungs has improved. Quantitative analysis of the chest CT scan shows 1% reticular opacity (orange color) and less than 1% honeycomb (red color).
GGA = ground glass attenuation

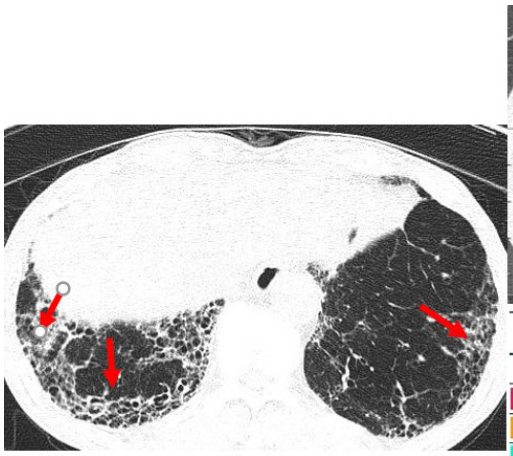


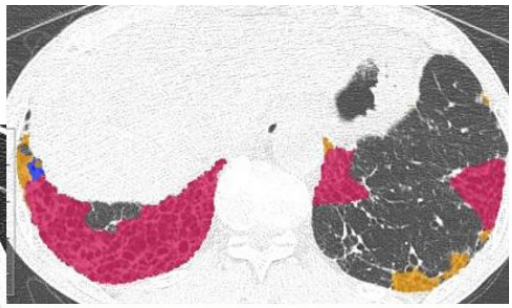
	Whole Lungs	Rt. Lung	Lt. Lung
Volume	3917 cc	2096 cc	1821 cc
Honeycomb	0 %	0 %	0 %
Reticular	2 %	2 %	2 %
GGO	0 %	0 %	0 %
Consolidation	0 %	0 %	0 %
Emphysema	0 %	0 %	0 %
Normal	98 %	98 %	98 %

	Whole Lungs	Rt. Lung	Lt. Lung
Volume	4039 cc	2148 cc	1892 cc
Honeycomb	0 %	0 %	0 %
Reticular	1 %	1 %	1 %
GGO	0 %	0 %	0 %
Consolidation	0 %	0 %	0 %
Emphysema	0 %	0 %	0 %
Normal	99 %	99 %	99 %

20, 21). 흉부 CT상 UIP 형태를 분할(segmentation)하는 AI 알고리즘은 주로 CNN을 이용하는데, 이 알고리즘은 정상인 폐 실질과 UIP 형태의 폐 실질을 구분하게 하는 기술이고, 연구자에 따라 정확도는 차이가 있지만 UIP 형태를 분할하는 정확성은 81.8%–98.5%로 높다(21). 그러나 아직까지 극복할 수 없는 UIP 형태를 분할하는 기술적인 오류는 주로 폐문부(hilum)와 기관지혈관가지(bronchovascular tree) 사이의 공간에서 생기며, 정상과 비슷한 음영을 보이는 UIP 형태, 폐기종이나 경화(consolidation)이 있는 경우에도 생긴다. CT상 UIP 형태를 분류(classification)하는 과정은 CT에서 보이는 소견을 정상, 간유리 음영, 망상형 음영, 벌집 모양, 폐기종, 경화로 분류를 하는 것이다(4, 5, 20-22). 이 과정은 분할 과정과 비교했을 때 정확성은 좀 더 떨어진다. Walsh 등(23)은 IPF가 있는 1157개의 CT를 흉부영상의학과 전문의와 AI로 UIP 진단 정확성을 비교했을 때

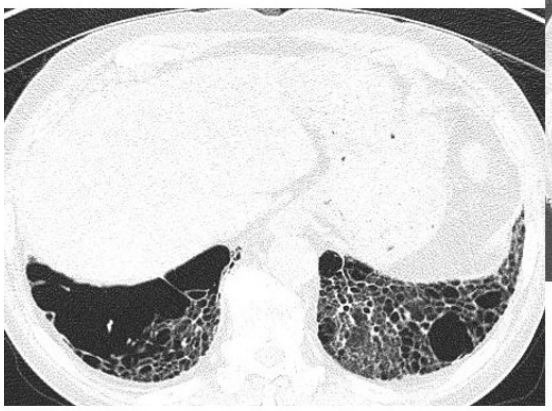
Fig. 7. A 74-year-old male patient current smoker (30 pack-years) with severe dyspnea underwent chest CT. Pulmonary function tests showed a reduced FEV1/FVC of 32% and a reduced DLCO of 4.8 mL/mmHg/min. Chest CT reveals honeycombing and reticular opacity with GGA in the subpleural area of both lungs. Quantitative analysis of the chest CT scan shows 4% reticular opacity (orange color) and 4% honeycomb (red color).
DLCO = diffusing capacity of the lung for CO, FEV1 = forced expiratory volume in one second, FVC = forced vital capacity, GGA = ground glass attenuation





	Whole Lungs	Rt. Lung	Lt. Lung
Volume	4981 cc	2395 cc	2586 cc
Honeycomb	4 %	5 %	2 %
Reticular	4 %	4 %	3 %
GGO	0 %	0 %	0 %
Consolidation	0 %	0 %	0 %
Emphysema	0 %	0 %	0 %
Normal	92 %	90 %	94 %

Fig. 8. A 78-year-old male patient current smoker (50 pack-years) with severe dyspnea underwent chest CT. Pulmonary function test showed a reduced FEV1/FVC of 73% and a DLCO of 7.9 mL/mmHg/min. Chest CT demonstrates honeycombing, reticular opacity, and lung destruction due to airway enlargement with fibrosis in the subpleural area of both lungs. Quantitative analysis of the chest CT scan shows 1% reticular opacity (orange color) and 5% honeycomb (red color). However, this software cannot measure lung destruction caused by airway enlargement with fibrosis and emphysema within the honeycombing areas.
DLCO = diffusing capacity of the lung for CO, FEV1 = forced expiratory volume in one second, FVC = forced vital capacity



	Whole Lungs	Rt. Lung	Lt. Lung
Volume	4374 cc	2365 cc	2009 cc
Honeycomb	5 %	2 %	8 %
Reticular	1 %	2 %	1 %
GGO	0 %	0 %	0 %
Consolidation	0 %	0 %	0 %
Emphysema	5 %	5 %	6 %
Normal	89 %	92 %	85 %

각각 70.7%와 73.3%를 보였고, Christe 등(24)도 307개의 CT로 흉부영상의학과 전문의와 AI로 UIP 진단 정확성을 비교했을 때 각각 80%와 79%를 보였다. 이런 연구를 통해서 CT에서 UIP 형태를 진단하는 AI 모델이 흉부영상의학과 전문의와 비슷한 정도의 진단을 보이는 것을 알 수 있다. 그러나 실제로 UIP를 진단하는 AI 모델의 정확성은 연구자마다 차이가 있다. 그 이유로는 UIP가 흔하지 않아서 데이터가 정상적으로 AI를 훈련하기에는 부족하거나 기술적으로 벌집 모양과 폐기종을 잘 구분하지 못하는 경우이다. 또한, UIP의 경우 폐 조직 검사를 대부분 하지 않기 때문에 확진을 영상의학과 전문의가 하는 데 이런 경우 AI를 훈련하고 테스트를 하기 위해 모은 데이터에서 확진에 대한 개인 간 편차가 발생할 수 있다. 이를 해결하기 위해서는 UIP에 대해 공신력 있는 대규모 공개 CT 데이터를 만들어서 UIP 진단 AI를 개발할 때 이용하도록 해야 한다.

마지막으로 흉부 CT상 UIP 형태의 정량화는 AI가 분류한 UIP 형태 소견들이 정상 폐 실질의 용량에 비해서 얼마나 차지하고 있는지 비율을 계산한다. 이 기술은 향후 UIP의 약물 치료에 대한 반응을 평가하는 데 유용하게 사용될 가능성이 있다(4, 20). 미국 식품의약국(United States FDA)에서는 이러한 방법을 지지해 왔고, 일부 폐 섬유화 치료제에 대해서는 AI를 이용한 정량화 방법으로 효과를 측정해 왔다. 예를 들면 IPF 환자에게 닌테다니브(nintedanib) 치료 효과에 대한 IIIb 임상 연구에서(phase IIIb trial)위약을 투여한 환자와 비교했을 때 흉부 CT상 폐 섬유화 점수(lung fibrosis score)가 더 적게 증가했다고 보고했다(20, 25). 물론 이런 AI가 여전히 기술적으로 한계가 있고, 외부 검증(external validation)이 필요하지만 미래에 흉부 CT에서 UIP 형태를 진단 후 모니터링하고, 약물 투여 후에 평가를 할 수 있는 영상 생물표지자(image biomarker)로써 역할을 할 수 있을 것으로 생각한다.

결론

실제로 흉부 CT에서 폐기종과 ILD에 대한 인공지능이 일상 업무에서 이 질환들을 해석하는 데 효율성과 정확성을 향상시켜 줌으로써 환자에게 개인화된 치료 계획을 세울 가능성이 있다. 그러나 그 결과를 최종적으로 해석하는 데는 여전히 숙련된 영상의학과 의사의 전문 지식과 임상 판단이 반드시 필요하다. 또한 흉부 CT상 폐기종과 ILD를 인공지능을 이용한 진단의 성능과 임상적 유용성을 향상시키기 위해 AI 알고리즘을 더욱 정교화하고 최적화하기 위한 지속적인 연구와 검증 연구가 필수적이다.

Author Contributions

Conceptualization, all authors; data curation, J.G.Y.; writing—original draft, all authors; and writing—review & editing, J.G.Y.

Conflicts of Interest

Gong Yong Jin has been a Section Editor of the Journal of the Korean Society of Radiology since 2024; however, he was not involved in the peer reviewer selection, evaluation, or decision process of this article. The remaining author has declared no conflicts of interest.

ORCID iDs

Sang Hyun Paik https://orcid.org/0000-0002-8276-333X
Gong Yong Jin https://orcid.org/0000-0002-1426-554X

Funding

None

REFERENCES

1. El Kaddouri B, Strand MJ, Baraghoshi D, Humphries SM, Charbonnier JP, van Rikxoort EM, et al. Fleischner society visual emphysema CT patterns help predict progression of emphysema in current and former smokers: results from the COPDGene study. *Radiology* 2021;298:441-449
2. Grenier PA. Deep learning assessment of emphysema progression at CT predicts outcomes. *Radiology* 2022;304:680-682
3. Humphries SM, Notary AM, Centeno JP, Strand MJ, Crapo JD, Silverman EK, et al. Deep learning enables automatic classification of emphysema pattern at CT. *Radiology* 2020;294:434-444
4. Dack E, Christe A, Fontanellaz M, Brigato L, Heverhagen JT, Peters AA, et al. Artificial intelligence and interstitial lung disease: diagnosis and prognosis. *Invest Radiol* 2023;58:602-609
5. Oh AS, Lynch DA, Swigris JJ, Baraghoshi D, Dyer DS, Hale VA, et al. Deep learning-based fibrosis extent on computed tomography predicts outcome of fibrosing interstitial lung disease independent of visually assessed computed tomography pattern. *Ann Am Thorac Soc* 2024;21:218-227
6. Al Wachami N, Guennouni M, Iderdar Y, Boumendil K, Arraji M, Mourajid Y, et al. Estimating the global prevalence of chronic obstructive pulmonary disease (COPD): a systematic review and meta-analysis. *BMC Public Health* 2024;24:297
7. Martini K, Frauenfelder T. Advances in imaging for lung emphysema. *Ann Transl Med* 2020;8:1467
8. Celli BR, MacNee W; ATS/ERS Task Force. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004;23:932-946
9. Lynch DA, Moore CM, Wilson C, Nevrekar D, Jennermann T, Humphries SM, et al. CT-based visual classification of emphysema: association with mortality in the COPDGene study. *Radiology* 2018;288:859-866
10. Šileikienė V, Urbonas M, Matačiūnas M, Norkūnienė J. Relationships between pulmonary function test parameters and quantitative computed tomography measurements of emphysema in subjects with chronic obstructive pulmonary disease. *Acta Med Litu* 2017;24:209-218
11. deSouza NM, Achten E, Alberich-Bayarri A, Bamberg F, Boellaard R, Clément O, et al. Validated imaging biomarkers as decision-making tools in clinical trials and routine practice: current status and recommendations from the EIBALL* subcommittee of the European Society of Radiology (ESR). *Insights Imaging* 2019;10:87
12. Messerli M, Ottilinger T, Warschkow R, Leschka S, Alkadhi H, Wildermuth S, et al. Emphysema quantification and lung volumetry in chest X-ray equivalent ultralow dose CT – Intra-individual comparison with standard dose CT. *Eur J Radiol* 2017;91:1-9
13. Smith BM, Austin JH, Newell JD Jr, D'Souza BM, Rozenshtein A, Hoffman EA, et al. Pulmonary emphysema subtypes on computed tomography: the MESA COPD study. *Am J Med* 2014;127:94.e7-94.e23
14. Lynch DA, Austin JH, Hogg JC, Grenier PA, Kauczor HU, Bankier AA, et al. CT-definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner society. *Radiology* 2015;277:192-205
15. Lee KS, Han J, Wada N, Hata A, Lee HY, Yi C, et al. Imaging of pulmonary fibrosis: an update, from the AJR special series on imaging of fibrosis. *AJR Am J Roentgenol* 2024;222:e2329119
16. Lynch DA, Sverzellati N, Travis WD, Brown KK, Colby TV, Galvin JR, et al. Diagnostic criteria for idiopathic pulmonary fibrosis: a Fleischner society white paper. *Lancet Respir Med* 2018;6:138-153
17. Byon JH, Jin GY, Han YM, Choi EJ, Chae KJ, Park EH. Quantitative CT analysis based on smoking habits and chronic obstructive pulmonary disease in patients with normal chest CT. *J Korean Soc Radiol* 2023;84:900-910
18. Chae KJ, Lim S, Seo JB, Hwang HJ, Choi H, Lynch D, et al. Interstitial lung abnormalities at CT in the Korean national lung cancer screening program: prevalence and deep learning-based texture analysis. *Radiology* 2023;307:e222828
19. Hida T, Nishino M, Hino T, Lu J, Putman RK, Gudmundsson EF, et al. Traction bronchiectasis/bronchiolecta-

sis is associated with interstitial lung abnormality mortality. *Eur J Radiol* 2020;129:109073
20. Walsh SLF, De Backer J, Prosch H, Langs G, Calandriello L, Cottin V, et al. Towards the adoption of quantitative computed tomography in the management of interstitial lung disease. *Eur Respir Rev* 2024;33:230055
21. Soffer S, Morgenthau AS, Shimon O, Barash Y, Konen E, Glicksberg BS, et al. Artificial intelligence for interstitial lung disease analysis on chest computed tomography: a systematic review. *Acad Radiol* 2022;29(Suppl 2):S226-S235
22. Huang S, Lee F, Miao R, Si Q, Lu C, Chen Q. A deep convolutional neural network architecture for interstitial lung disease pattern classification. *Med Biol Eng Comput* 2020;58:725-737
23. Walsh SLF, Calandriello L, Silva M, Sverzellati N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. *Lancet Respir Med* 2018;6:837-845
24. Christe A, Peters AA, Drakopoulos D, Heverhagen JT, Geiser T, Stathopoulou T, et al. Computer-aided diagnosis of pulmonary fibrosis using deep learning and CT images. *Invest Radiol* 2019;54:627-632
25. Lancaster L, Goldin J, Trampisch M, Kim GH, Ilowite J, Homik L, et al. Effects of nintedanib on quantitative lung fibrosis score in idiopathic pulmonary fibrosis. *Open Respir Med J* 2020;14:22-31

폐기종 및 간질성 폐질환: 인공지능 소프트웨어 사용 경험

백상현[1] · 진공용[2]*

흉부 CT상 폐기종이나 간질성 폐질환의 형태나 범위를 인공지능을 이용하여 자동적으로 객관적으로 진단하는 다양한 알고리즘을 개발되고, 이를 증명하는 연구들이 진행되어 왔다. 흉부 CT상 인공지능을 이용한 폐기종 정량화 연구들을 보면 CT상 폐기종의 상대적인 양이 증가와 폐 기능의 악화와 연관이 있으며, 특히 중심성 폐기종을 중심으로 정량화를 하는 것이 임상 증상이나 만성폐쇄성 폐질환의 사망률을 예측하는 데 도움이 된다고 보고하고 있다. 또한, 간질성 폐질환에서는 인공지능이 CT상 통상성 간질성 폐렴의 형태를 정상, 간유리 음영, 망상형 음영, 벌집 모양, 폐기종, 경화로 분류를 할 수 있고, 인공지능이 흉부영상의학과 전문의와 비슷한 정도로 통상성 간질성 폐렴을 진단(70%–80%) 할 수 있다고 보고했다. 그러나 인공지능의 결과들이 흉부 CT의 스캔 변수들, 재구성 알고리즘, 방사선 선량, 개발된 인공지능 훈련 데이터에 의해 영향을 받으며, 이러한 이유로 아직까지 흉부 CT상 폐기종과 간질성 폐질환의 진단과 정량화는 실제로 일상 업무에서 제한적으로 사용되고 있다. 이 논문에서는 폐기종과 간질성 폐질환의 진단과 정량화를 위해서 인공지능을 사용하고 있는 저자들의 경험을 증례로 소개를 하고, 이 두 질환의 인공지능의 효용성과 제한점에 대해서 언급하고자 한다.

[1]한국영상의원 흉부 분야,
[2]전북대학교 의과대학 전북대학교병원 임상의학연구소-의생명연구원, 영상의학과