Review Article

J Korean Soc Radiol 2022;83(3):473-485 https://doi.org/10.3348/jksr.2022.0048 eISSN 2288-2928

Brain MRI-Based Artificial Intelligence Software in Patients with Neurodegenerative Diseases: Current Status 퇴행성 뇌질환에서 뇌 자기공명영상 기반 인공지능 소프트웨어 활용의 현재

So Yeong Jeong, MD¹ (D), Chong Hyun Suh, MD^{2*} (D), Ho Young Park, MD² (D), Hwon Heo, PhD² (D), Woo Hyun Shim, PhD² (D), Sang Joon Kim, MD² (D)

¹Department of Radiology, Hanyang University Medical Center, Hanyang University Medical College, Seoul, Korea

²Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

The incidence of neurodegenerative diseases in the older population has increased in recent years. A considerable number of studies have been performed to characterize these diseases. Imaging analysis is an important biomarker for the diagnosis of neurodegenerative disease. Objective and reliable assessment and precise detection are important for the early diagnosis of neurodegenerative diseases. Artificial intelligence (AI) using brain MRI applied to the study of neurodegenerative diseases could promote early diagnosis and optimal decisions for treatment plans. MRI-based AI software have been developed and studied worldwide. Representatively, there are MRI-based volumetry and segmentation software. In this review, we present the development process of brain volumetry analysis software in neurodegenerative diseases, currently used and developed AI software for neurodegenerative disease in the Republic of Korea, probable uses of AI in the future, and AI software limitations.

Index terms Neurodegenerative Disease; Brain; Magnetic Resonance Imaging; Artificial Intelligence

대한영상의학회지

Received April 8, 2022 Revised May 5, 2022 Accepted May 15, 2022

*Corresponding author

Chong Hyun Suh, MD Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea.

Tel 82-2-3010-1648 Fax 82-2-476-4719 E-mail chonghyunsuh@amc.seoul.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.

ORCID iDs

So Yeong Jeong 🕩 https:// orcid.org/0000-0003-4705-0008 Chong Hyun Suh 匝 https:// orcid.org/0000-0002-4737-0530 Ho Young Park 🝺 https:// orcid.org/0000-0003-2318-9806 Hwon Heo 匝 https:// orcid.org/0000-0002-6103-4680 Woo Hyun Shim 匝 https:// orcid.org/0000-0002-7251-2916 Sang Joon Kim 厄 https:// orcid.org/0000-0001-7070-7333

서론

현대사회가 점차 고령화 사회가 됨에 따라 알츠하이머병, 파킨슨병 등과 같은 퇴행성 뇌

질환의 발병률이 증가하고 있으며, 그 사회적 중요성이 커지고 있다. 퇴행성 뇌질환에서 뇌 자기 공명영상(이하 MRI)을 이용한 구조적 이상과 뇌의 위축 정도 및 위축 패턴의 평가는 퇴행성 뇌질 환을 진단하는 하나의 영상표지자이다(1, 2). 하지만 임상현장에서는 뇌의 위축의 정도와 패턴의 평가가 영상의학과 전문의의 경험과 시각적 평가에 주로 의존하고 있어, 주관적인 평가가 되고 있 다고 할 수 있다(3).

최근 다양한 임상용 뇌용적 정량 분석 소프트웨어들이 소개되고 있고 뇌 MRI를 이용한 뇌용적 분석 소프트웨어의 임상적 적용에 대한 전문가 의견과 권고안에서 2020년 8월까지 국내 승인된 소프트웨어를 소개한 바 있다(4). 이 소프트웨어들은 뇌의 회색질 및 백질의 부피를 자동 정량 분 석하여, 특정 뇌영역의 위축 정도를 평가하고, 내장되어 있는 연령별정상규준 자료(normative data)를 바탕으로 연령별 정상규준 백분위수(normative percentile)와 같은 통계수치를 제공하 여 주관적 평가의 한계점을 어느 정도 극복할 수 있게 되었다(5-7). 임상용 정량 분석 소프트웨어 들은 3D T1 강조영상을 바탕으로 1.5T 이상의 MRI 기기에서 5-6분 내로 결과를 얻을 수 있어, 다 수의 의료기관에서 비교적 쉽게 활용할 수 있다. 현재 다양한 임상용 소프트웨어들이 미국과 유럽 식품의약안전청에서 임상용으로 승인을 받아, 정식으로 보험청구 가능 항목으로 인정을 받고 있 으며, 국내에서도 식품의약품안전처의 승인을 받아 여러 병원에서 진료에 이용되고 있다.

초기의 수동 혹은 반 자동화 정량 분석 소프트웨어는 사용자 인터페이스가 어렵고, 표준화의 문 제, 긴 프로세싱 시간과 같은 문제점으로 실제 임상진료에 사용되기에는 한계가 있었다. 하지만 이후에 이러한 한계점을 극복한 완전 자동화 뇌용적 정량 분석 소프트웨어가 개발되었고, 현재 실 제 임상에서 사용되기 시작했다(3, 4). 과거 주관적 판단과 수동 혹은 반자동 연구용 정량 분석 소 프트웨어의 단점을 극복한 외국에서 개발된 완전 자동화 임상용, 뇌용적 정량 분석 소프트웨어들 은 주로 서양인 데이터를 기반으로 만들어져서 한국인 데이터 기반 소프트웨어 개발의 필요성이 대두되었다. 이에 2017년부터 국내에서도 인공지능을 이용한 정량적 분석 및 분류 임상용소프트 웨어가 개발되었으며(4), 2020년 후반기 이후 양적 질적 성장이 급격해지고 있다. 따라서, 최근 2 년여간의 소프트웨어 개발과 이용현황을 영상의학과 전문의들이 공유할 필요가 있다.

본 종설에서는 퇴행성 뇌질환에서 정량적 분석 소프트웨어의 발전과정을 소개하고, 현재 국내 에서 개발된 인공지능 소프트웨어들이 어떻게 이용되고 있는지, 또 앞으로 어떻게 인공지능을 이 용할 수 있는지와 인공지능 소프트웨어의 제한점에 대해 소개하고자 한다.

뇌용적 분석 소프트웨어의 발전

뇌 부피의 시각적 평가: 내측 측두엽 위축 척도(Medial Temporal Lobe Atrophy Scale)

내측 측두엽 위축변화는 알츠하이머병의 아밀로이드병리, 타우병리, 신경퇴행병리 등 3가지 주 요 병리 중 한 축을 이루는 소견으로, 알츠하이머성 치매의 진단에 오랫동안 중요하게 활용되었 다. MRI를 이용한 뇌위축 평가는 1992년 Scheltens 등(8)에 의해 제안된 내측 측두엽 위축 정도를 등급화한 Scheltens 척도가 널리 쓰이고 있다. 내측 측두엽위축 등급의 알츠하이머성 치매 진단능 을 살펴본 21개의 연구를 기반으로한 메타분석에서 정상군 대비 민감도 74%, 특이도 88%, area

under the curve (이하 AUC) 0.88로 보고되었고 내측 측두엽위축 척도와 해마 용적측정의 진단 능비교시 통계적으로 유의한 차이는 없었다(*p* = 0.04) (9). 반면, 실제 진료에서는 관찰자 간 일치 도(interobserver agreement)는 변동폭이 크며 일부 연구에서는 전문가와 비전문가 간의 차이도 문제로 지적되었다(3, 10). 위의 메타분석에서, 관찰자 간 일치도와 관측자 내 일치도(intraobserver agreement)는 상관계수(correlation coefficient)가 각각 0.83, 0.85로 높은 편이었지만, 이러한 작은 차이도 추적검사를 포함하여 많은 영상을 보게 되면 차이가 누적될 우려가 있다(9).

자동 뇌용적 분석 소프트웨어의 개발

이후 반자동 혹은 자동화 분할에 의한 해마 용적측정의에 대한 연구가 많이 되었고 33개 논문을 기반으로한 메타분석에서 해마 부피를 이용한 정상인 대비 알츠하이머성 치매 진단능이 민감도 82%, 특이도 87%, AUC 0.91로 보고되었다(11). 하지만 연구용뇌용적 분석 소프트웨어는 표준화 문제, 복잡한 사용자 환경, 긴 프로세싱 시간 등의 문제가 있었고(12), 이런 문제점을 극복한 여러 완전 자동화 뇌용적 분석 임상용 소프트웨어들이 미국 식품의약안전청 및 유럽의 승인을 받고 사용 중에 있다. 이 중에서 최초로 개발되고 승인된 Cortechs사의 NeuroQuant[®]가 현재 가장 많이 사용되고 있으며, 연구가 많이 된 소프트웨어 중 하나로 검증이 가장 많이 되었다(Table 1). NeuroQuant[®]는 프로세싱 시간이 10분 정도로 기존 연구용 소프트웨어대비 획기적으로 시간을 단축 하였고, 전체 뇌용적을 보정한 각 영역별 부피와 연령에 따른 정상규준 백분위수(normative percentile)를 제공한다. 절대표준 혹은 상대표준으로 간주되는 기존 연구용 소프트웨어와의 일치도 가 비교적 높으며(5, 6, 13-15), MRI 기반의 알츠하이머성 치매 진단능도 AUC 0.69-0.88로 보고하고 있다(Table 1) (6, 14, 16, 17).

알츠하이머병에서 인공지능 기반 뇌용적 분석 소프트웨어의 개발

세계적으로 딥러닝을 바탕으로 한 인공지능 연구에서 여러 방법들 중 합성곱 신경망(convolutional neural network; 이하 CNN)을 이용하여 우수한 연구 결과를 보였고, 인공지능을 기반으로 하는 여러 소프트웨어들에서 CNN을 주요 기법으로 한 소프트웨어들이 다수를 이루고 있다(18).

현재까지 국내외에서 의료영상처리 기반 자동 뇌용적 분석 소프트웨어 중 미국(Food and Drug Administration), 유럽(Conformite Europeenne Mark), 또는 한국 식품의약품안전처에서 승인 한 임상 적용이 가능한 인공지능기반 소프트웨어는 Table 2 (국내)와 Table 3 (국외)과 같다. 이들 은 딥러닝을 포함한 인공지능기법을 현재 대부분 적용하고 있다.

한국의 체외진단의료기기 등급분류는 체외진단의료기기법 시행규칙에 따라 4개 등급으로 분류 한다(19). 1등급 체외진단의료기기는 개인과 공중보건에 미치는 잠재적 위해성이 낮은 기기로 신 고 절차가 필요하며, 2등급 기기는 개인에게 중증도의 잠재적 위해성을 가지며 공중보건에 미치 는 잠재적 위해성이 낮은 기기로 한국의료기기안전정보원 인증이 필요하다. 3등급 기기는 개인에 게 고도의 잠재적 위해성을 가지며 공중보건에 중증도의 잠재적 위해성을 가진 기기이며, 4등급 기기는 개인과 공중보건에 고도의 위해성을 가진 기기로, 3·4등급 기기는 식품의약품안전처 허가 절차가 요구된다.

$VeuroQuant^{\mathbb{R}}$
÷
0
Performance
<u>.</u> 2
iagnost
anc
Correlation
ie Correlatior
the Correlation
on the Correlation
Research on the Correlatior
 Research on the Correlatior

					Correl	ation	Dia	ignostic Perfor	mance	
Author	Year	Journal	Population	Control	Comparison	Coefficient	Method	Sensitivity (%)	Specificity (%)	AUC
Brewer et al. (5)	2009	AJNR Am J Neuroradiol	20 mild probable AD	20 healthy	Manual segmentation	ICC, 0.61–0.99				
Ochs et al. (13)	2015	J Neuroimaging	20 AD	20 healthy	FreeSurfer	ICC, 0.62–0.99				
Tanpitukpongse et al. (6)	2017	AJNR Am J Neuroradiol	85 MCI to AD converter	107 MCI nonconverter	Neuroreader	r, 0.60–0.99	Hippocampus			0.69
Min et al. (17)	2017	AJR Am J Roentgenol	30 mild-to- moderate AD	25 healthy			MTL index A	63.3	100	0.83
Persson et al. (14)	2018	Acta Neurol Scand	31 AD	14 SCI, 11 MCI	MTA score	r, -0.77	Hippocampus	71	92	0.88
Lee et al. (15)	2021	Korean J Radiol	42 AD, 85 MCI	45 healthy	Inbrain	r, 0.72–96 ICC, 0.83–0.98				
Kim et al. (16)	2021	Sci Rep	26 AD	55 MCI, 32 healthy			Hippocampus	54	76	0.78
AD = Alzheimer's d	isease,	AUC = area under the curve,	ICC = intraclass corre	elation, MCI = mild e	cognitive impairme	ent, MTL = medial	temporal lobe, SC	CI = subjective c	cognitive impa	irment

현재 우리나라에서 개발되고 인증된 임상용 소프트웨어들을 살펴보면 다음과 같다. 2017년 4월 에 마이다스 아이티의 Inbrain[®] 소프트웨어가 국내 최초 개발 및 인증을 받았으며, 삼성서울병원 데이터를 바탕으로 개발되었고 FreeSurfer 플랫폼 기반으로 뇌백질 분획화와 skull stripping 과

Table 2. Artificial Intelligence-Augmented Automated Brain Volumetry and Classification Software in the Republic of Korea

Year	Korean MFDS	Software	Company	Key Features	Main Techniques	Data	Evidence
2017-04	Grade 2	Inbrain Morph	MIDAS IT	Brain volumetry	CNN (mainly for skull stripping and white matter segmentation)	Samsung Medical Center (3000 >)	Korean J Radiol 2021 (15), J Clin Neurol 2021 (53)
2018-08	Grade 2	Neuro I	NEUROZEN	Brain volumetry	Atlas-based	National Research Center for Dementia (2000 >)	
2019-06	Grade 2	DeepBrainTM	VUNO	Brain volumetry	CNN (2.5D HighResNet) + XGBoost	Asan Medical Center	AJNR Am J Neuroradiol 2020 (20)
2019-08	Grade 2	ASTROSCAN	JLK	Brain volumetry	3D U-Net		
2020-12	Grade 3	DeepBrain AD	VUNO	Risk classification of AD	CNN (Inception v4)	Seoul National University Bundang Hospital	Sci Rep 2020 (21)
2021-03	Grade 3	Neurophet AQUA 2.0	Neurophet	Brain volumetry	U-Net++ (ResNet with split-attention)	Yeouido St. Mary's Hospital, ST. Vincent's Hospital	Sci Rep 2020 (21), Brain Sci 2021 (24), Diagnostics (Basel) 2020 (23)

AD = Alzheimer's disease, CNN = convolutional neural network, MFDS = Ministry of Food and Drug Safety, NA = nonavailable

Table 3. Artificial Intelligence-Augmented Automated Brain Volumetry or Classification Software from Global Companies (Except the Republic of Korea)

Software	Company	Country	Key Features	Main Techniques	CE Mark	US FDA
AIRAscore structure	AIRAmed	Germany	Brain volumetry	NA	Class I	
Diadem	Brainminer	UK	Brain volumetry	Multi-atlas/MRF	Class I	
Neuroreader [®]	Brainreader	Denmark	Brain volumetry	Multi-atlas	Class I	Class II
cNeuro [®] cDSI	Combinostics	Finland	Clinical decision support in dementia	MRF (graph-cut)/EM	Class IIa	Class II
NeuroQuant [®]	Cortechs.ai	Netherland	Brain volumetry	Multi-atlas (dynamic atlas)	Class IIa	Class II
THINQ	CorticoMetrics	US	Brain volumetry	Atlas-based/MRF (FreeSurfer)	Class II
icobrain DM	icometrix	Belgium	Brain volumetry	Multi-atlas	Class I	Class II
mdbrain	Mediaire	Germany	Brain volumetry	U-Net (ensemble)	Class I	
Pixyl.Neuro.BV	Pixyl	France	Brain volumetry	MRF/EM	Class II	
Quantib [®] ND	Quantib	Netherland	Brain volumetry	NA	Class IIa	Class II
Neurocloud VOL	Qubiotech	Spain	Brain volumetry	NA	Class I	
Brain atrophy screening	Quibim	Spain	Brain volumetry	CNN (2.5 D)	Class II	
QyScore	Qynapse	France	Brain volumetry	NA	Class II	Class II
Al-Rad Companion Brain MR	Siemens Healthineers	Germany	Brain volumetry	MRF/EM (tissue) Multi-atlas (region)	Class IIa	Class II

CE = Conformite Europeenne, CNN = convolutional neural network, EM = expectation-maximization algorithm, MRF = markov random field, NA = nonavailable



2018년 8월 뉴로젠(Neurozen)의 뉴로아이가 의료영상분석장치 소프트웨어 2등급 제품으로 식 약처 승인을 받았다. 뉴로아이는 아틀라스 기반 뇌용적 분석 소프트웨어로 광주치매코호트연구단 데이터를 바탕으로 개발되었고 Inbrain[®]과 비슷한 기능을 제공한다. 주로 코호트연구단 데이터 를 분석하는데 이용되고 있으며, 타 소프트웨어와의 비교 검증연구는 아직 없다.

2019년 6월 뷰노의 DeepBrain이 식약처 승인(2등급)을 받았다. 뷰노의 DeepBrain은 서울아산 병원과 뷰노가 공동으로 개발하였고 강동경희대병원 데이터와 공적 데이터베이스인 Alzheimer's disease neuroimaging initiative (ADNI)와 Open Access Series of Imaging Studies (OASIS)의 데이터를 이용하여 외부 검증하였다(20). 모델은 2.5D ResNet 기반 뇌분할 알고리즘으로 44개의 convolutional layer를 갖는 2.5 channel high Resnet architecture를 이용하여 104개의 뇌 영역 으로 나누고 뇌분할에 걸리는 프로세싱 시간은 1분 정도이다. 뇌용적뿐 아니라, 대뇌 피질(cortical thickness) 정보, 뇌백질 고신호강도 병변의 부피 정보도 제공하고 있다(Supplementary Figs. 1-3 in the online-only Data Supplement).

이후 제이엘케이(JLK)의 ASTROSCAN도 3D U-Net 모델을 이용한 뇌용적 분석 소프트웨어로 식약처 승인(2등급)을 받았다. ASTROSCAN는 대뇌 회피질 영역을 62가지 영역으로 분류하고, 각 영역별 대뇌 피질 두께를 측정하여 정상, 경도인지장애, 알츠하이머성 치매에 대한 확률 정보를 제공한다. 공개 데이터를 이용한 뇌분할 성능 분석에서 다이스계수 0.86을 보였고 분석시간 20분 이내로 보고 하고 있다. 아직 임상경험 및 검증연구는 없다.

2020년 12월, 뷰노의 DeepBrain AD는 Inception-v4 기반으로 알츠하이머성 치매 위험도를 제 시하는 소프트웨어로 식약처 승인받아 퇴행성 뇌질환 분야에서는 최초로 의료영상분석장치 소프 트웨어 3등급 제품이 되었다. DeepBrain AD는 분당서울대병원과 공동개발하였고, Inception-v4 architecture를 이용하여 알고리즘을 구성하였다(21). 알츠하이머성 치매에서 특이적으로 나타나 는 공간적인 뇌 위축 패턴과 질감을 학습한 인공지능 기반 비정상 판별모델을 이용하여 자동으로 알츠하이머성 치매의 가능성을 수치화하여 제공하고 가능성에 따라 알츠하이머성 치매 예측 지 수를 0에서 100점으로 표시한다. 평균 프로세싱 시간은 23-24초이고, 성능은 알츠하이머성 치매 와 정상인 감별 기준 외부 데이터셋에서 AUC 0.88-0.89 정도가 나왔다(21). DeepBrain AD는 현 재의 생체표지자기준으로 정의된 알츠하이머병을 완전히 반영하지는 않는다.

2021년 3월에는 뉴로핏이 U-Net 기반의 아쿠아 2.0를 개발해 식약처 승인(2등급)을 받았다. 아 쿠아 2.0은 여의도성모병원과 성빈센트병원의 데이터를 이용하여 개발되었고 모델은 3차원 이미 지 패치 기반의 Split-Attention U-Net을 이용하였다(22). 아쿠아 2.0은 뇌 97개 영역으로 구획화하 고 영역별 구조 정보를 분석해 퇴행성 뇌질환과 관련한 비정상적인 위축, 구조 변화와 해부학적 비대칭성, 백질병변을 분석하며, 프로세싱은 1분여이다(Supplementary Fig. 4 in the online-only Data Supplement) (21, 23, 24). 현재 이용 경험이나 검증 연구는 아직 없다.

대한영상의학회지

혈관성 치매 혹은 치매의 혈관성병변에 대한 인공지능 적용

혈관성 치매는 우리나라에서 두 번째로 흔한 치매의 원인으로 치매의 약 10%-20%의 원인으로 보고되고 있다(25). 현재 혈관성 치매는 NINDS-AIREN criteria의 기준을 따르며, 이에 따르면 반 드시 뇌백질병변이 MRI에서 확인되어야 한다(26). 따라서 백질의 고신호 병변을 정확하게 평가하 는 것이 혈관성 치매의 진단에 중요하다. 백질의 고신호 병변은 Fazekas 척도를 이용하여 평가하 며, 뇌실 주변부 백질(periventricular white matter)와 심부 백질(deep white matter)로 나누어 서 각각 0에서부터 3까지 4단계로 구분한다(27). Fazekas 척도는 관찰자의 주관적 평가에 의존적 이며, 시간에 따른 변화를 분석하는 것이 어렵다는 단점이 있어, 국내 연구자들은 좀 더 재현성이 높은 visual scale을 보고하기도 하였다(28).

따라서 이를 극복하기 위해 완전 자동화 분할 소프트웨어가 개발되었고, 대표적으로 Lesion-Quant[®]가 있다. 소프트웨어는 백실 고신호 병변을 구획화하고 부피를 정량화하며, 시간에 따른 부피의 변화를 평가할 수 있다. 그 외 국내에서 개발한 대부분의 소프트웨어는 딥러닝 기반으로 수초 내에 대뇌 백질 고신호 강도를 자동으로 분할해 평가하며 백질 고신호 영역의 부피와 전체 백질의 부피의 비율을 리포트에 같이 보고하고 있다(Supplementary Fig. 3 in the online-only Data Supplement).

파킨슨병에서의 인공지능 소프트웨어의 적용

파킨슨 병 진단에 있어 MRI 영상을 이용한 나이그로좀(nigrosome) 영상표지자는 매우 유용하 다고 알려져 있다(29). 파킨슨병에서는 흑질(substantia nigra)에서 나이그로좀이 소실되는 것이 병리 특징이며(30, 31), 특히 나이그로좀-1 소실이 가장 먼저 일어나기 때문에 흑질에서 나이그로 좀을 1부터 5까지 세분화하여 나이그로좀-1의 소실을 찾는 것이 파킨슨병의 초기진단에 유용하다 고 알려져 왔다(32). 나이그로좀-1 소실의 진단능을 19개의 논문으로 분석한 메타분석에서 파킨슨 병 또는 원발성 파킨슨 증후군 환자를 정상인과 구분함에 있어 자화율 강조 자기공명영상(susceptibility-weighted imaging; 이하 SWI)의 진단능을 조사한 결과, 파킨슨병/정상인에서 민감도 96%, 특이도 95%, 원발성파킨슨 증후군/정상인에서 민감도 87%, 특이도 93%를 보였다(33). 하지 만 나이그로좀-1 소실 소견은 MRI 영상품질에 따라 영향을 많이 받고, MRI에서 비진단율(nondiagnostic scan)이 4.2%로 보고되었다(33). 또한 나이그로좀-1 소실의 판단은 영상의학과 전문의 의 주관적인 평가를 바탕으로 하는 단점이 있다.

초기 SWI는 현재 파킨슨병 진단의 절대표준인 FP-CIT PET보다 진단능이 떨어지며, 평가의 객 관화가 쉽지 않다는 한계점이 있었다. 이를 개선하기 위하여, 대조도를 증가시킨, susceptibility map-weighted imaging (이하 SMwI)가 개발되었다(34). 하지만 SMwI는 MATLAB-based tool 을 이용한 후처리가 필요하고 정성적 분석이라는 한계가 있으며, reconstruction 방향도 일반적 인 axial image가 아닌 익숙하지 않은 oblique coronal image를 이용하여 영상 해석이 어려워 판 독자의 숙련도에 좌우된다는 한계가 존재했다.

이에 인공지능을 이용한 MRI 영상의 전처리 개선기술 및 나이그로좀 영역 자동 검출기술로 흑



퇴행성 뇌질환 분야에서 인공지능 적용의 미충족 수요

비전형적 파킨슨 증후군

특발성 파킨슨병(idiopathic Parkinson disease)과 진행성핵상마비(progressive supranuclear palsy; 이하 PSP)와 다계통위축(multiple system atrophy; MSA)과 같은 비전형적 파킨슨 증후 군을 구분하는 것은 임상적 진단기준에 의존하고 있다(36-38). 하지만 비교적 질환의 초기 단계에 서 임상적 진단기준에 의존하여 진단하는 것은 쉽지 않다. 각각의 질환들은 질환의 특징에 따라 뇌간의 각각 세부 영역 중 손상받는 부분이 달라 위축 정도가 다르게 나타난다. 따라서 이를 이용 하여 MRI를 이용한 파킨슨병을 진단하기 위한 여러 지표들이 연구되어왔고, 대표적으로 MR parkinsonism index가 있다(39). MR parkinsonism index에 관한 14개의 논문을 분석한 메타분 석 연구 결과를 보면 MR parkinsonism index는 PSP를 파킨슨병과 구분하는데 있어 민감도 96%, 특이도 98%, AUC 0.99을 보였다(33). 앞서 설명한 알츠하이머병의 진단을 위해 개발된 완전 자동화 뇌용적 분석 소프트웨어와 인공지능 기반 뇌용적 분석 소프트웨어들은 소뇌와 뇌간의 용 적값을 결과값으로 나타내지만, 중간뇌(midbrain), 뇌교(pons), 중소뇌각(middle cerebellar peduncle)과 같은 뇌간 각각의 세부 영역의 용적을 구분하여 나타내지는 못한다. 뇌간의 세부 용적 분석을 분석하는 것이 비전형 파킨슨 증후군과 초기 단계의 파킨슨병을 진단하는데 중요하고, 머 신 러닝을 이용한 뇌간의 세부 용적분석을 이용한 소프트웨어는 진단능이 AUC 0.77-0.98으로 보 고되었다(40). 인공지능 소프트웨어에도 이러한 뇌간의 세부 영역 구획화 및 용적 결과를 제시되 길 기대한다.

특발성 정상압 수두증

특발성 정상압 수두증은 교정 가능한 치매의 원인 중 하나로, 치매의 5%-6%의 원인으로 보고

되고 있고, 노령 인구에서 유병률이 0.5%-3%까지로 진단이 매우 중요하다고 할 수 있다(41-43). 특발성 정상압 수두증은 다른 퇴행성 뇌질환들과 비슷한 임상 양상을 보이거나 증상이 모호하게 나타나는 경우가 많아 임상 양상만으로 특발성 정상압 수두증을 진단하는 것은 쉽지 않으며, 영상 의학적 소견을 통해 진단에 도움을 받을 수 있다(44). 특발성 정상압 수두증의 특징적 이미지 소견 으로 2012년 일본의 가이드라인에서 diappropriately enlarged subarachnoid space hydrocephalus (이하 DESH)를 제시하였고, 예각 이하의 뇌량 각도(callosal angle)도 널리 쓰이고 있다 (43, 45). 10개의 논문을 기반으로한 메타분석에서 뇌량 각도의 정상압 수두증 진단능은 민감도 91%, 특이도 93%로 비교적 높게 보고되었다(46), 하지만 다른 10개의 논문을 기반으로한 메타분 석에서 DESH 소견은 특발성 정상압 수두증에서 44%에서만 보였다(47). 다른 연구에서도 DESH 의 민감도가 50% 미만으로 높지 않으며 일치도도 높지 않다는 한계가 지적되었다(48). 특발성 정 상압수두증의 영상 소견들은 평가자의 주관적 평가에 의존하며, 개개인에 따른 측정의 오차도 있 을 수 있어, 좀 더 정확하고 객관적인 평가가 요구된다. 인공지능을 이용하여 비조영증강 CT 이미 지에서 뇌실을 구획화하고, 수두증과 정상을 구분하도록 학습시켰을 때, 인공지능을 이용한 모델 은 민감도 93.6%, 특이도 94.4%, 및 AUC 0.93의 성능을 보였고, 이 모델의 정확도는 영상의학과 전문의의 정확도와 대등하였다(49). 그 외, 머신러닝을 이용하여 DESH 소견을 판단하는 모델에 대한 연구에서 전문가와 의견 일치 정도가 AUC 0.99를 보였다(50).

현재 대부분의 인공지능 소프트웨어에서 뇌실에 대한 부피 정보를 제시하고 있어 수두증 여부 판단이나 추적검사에서 도움이 될 것으로 판단된다. 하지만 아직 수두증과 관련한 인공지능의 연 구는 초기 단계이고 추가적인 연구가 필요할 것으로 생각된다.

인공지능 소프트웨어의 제한점 및 제언

현재 퇴행성 뇌질환에서 인공지능 소프트웨어의 역할은 아직까지 진입 단계로 생각되며, 대표 적인 제한점은 다음과 같다. 첫 번째, 인공지능 기반 뇌 자동분할 기술은 비교적 잘 작동하고 있지 만 아직 충분히 검증되지 않았다. 기존에 사용되던 뇌용적 소프트웨어들과의 뇌분할 성능 비교 연 구가 충분히 시도되어야 할 것으로 보인다. 두 번째, 인공지능 기반의 치매 분류 기술은 알츠하이 머성 치매와 다른 치매와의 구분은 어려운 실정으로 임상적용이 아직은 쉽지 않다. 현재까지 대부 분 인공지능 모델의 검증은 퇴행성 뇌질환 환자군과 정상군을 구분하는데 치중하였다면, 실제 의 료환경에서는 인지 저하를 호소하는 다양한 환자군 속에서 특정 질환을 진단하는 능력이 중요하 기 때문이다. 셋째, 젊은 연령이나 90세 이상의 연령대의 정상 MRI 데이터가 부족하여 해당 연령 대의 정상 표준치 확립이 어렵다. 정상 표준치 확립을 위한 전문가의 공동의 노력이 필요하다(4). 마지막으로, 퇴행성신경질환 진단에서 MRI의 역할이 특정 질환의 진단에 필요한 여러 생체표지 자 중 단 하나의 표지자에 해당한다는 점도 MRI 기반 인공지능 소프트웨어 기반 진단 및 평가에 서 반드시 고려해야 할 점이다.

의료영상을 이용한 진단 분석에서 인공지능을 이용한 연구들을 분석한 결과 외부 데이터 셋을 이용한 검증이 이루어진 경우는 6% 정도로, 아직까지 인공지능 기반의 소프트웨어들의 소프트웨



어 간 검증 혹은 임상검증 연구가 부족한 실정이다(51). 특히 우리나라에서 이미 상용화된 소프트 웨어조차 논문으로 보고된 경우는 Inbrain[®], DeepBrain, DeepBrain AD, AQUA 2.0에 불과하고 이 또한 검증이 불충분하다(Table 2). 또한 대부분의 소프트웨어 개발 업체들은 개발 시 기반한 인 구통계 자료(환자군의 수, 나이, 성별, 인종, 지역, 모집한 센터, 포함 혹은 제외 기준 등)와 MRI 데 이터에 대한 기본 자료(MRI vendor, MRI parameter 등)를 완전히 공개하지 않고 있다. 소프트 웨어 개발이나 소프트웨어 업그레이드 시 개발에 기반한 데이터에 대한 정보와 데이터 공개 기준 마련이 필요할 것으로 판단된다.

비교적 많은 검증이 된 NeuroQuant[®]와 같이(Table 1) 인공지능 소프트웨어들이 좀 더 정교하 고 업그레이드되고, 실용적으로 이용되기 위해서는 다양한 기관에서 충분한 검증이 이루어져야 하고 충분한 정보가 공개되어야 할 것이다. 이와 더불어 정상 표준치 확립, 프로토콜 및 시퀀스의 표준화, 분석 전후의 전문가 모니터링, 영상 획득, 분석, 판독 최적화를 위한 전문의의 노력이 권고 되고 있다(4). 결국 앞으로 인공지능 기반의 연구 및 소프트웨어 개발이 더 발전하고, 실제 환자 진 료에 실용적으로 적용되기 위해서는 무엇보다도 신뢰할 수 있는 방법을 이용한 성능 검증의 과정 이 필요할 것으로 판단된다(51, 52).

결론

퇴행성 뇌질환은 현재 인공지능을 이용하여 활발히 연구되고 있는 분야이며, 인공지능 기반 소 프트웨어들이 개발되어 실제 임상 진료에 도움이 되는 부분이 증가하고 있다. 향후 이러한 인공지 능 소프트웨어들의 다양한 어플리케이션들이 추가로 개발되어 정밀 의료에 도움이 될 것으로 기 대된다. 이를 위해, 인공지능 소프트웨어들의 추가적인 검증 및 전문가 집단의 다각도의 노력이 필요할 것으로 사료된다.

Supplementary Materials

The online-only Data Supplement is available with this article at http://doi.org/10.3348/jksr.2022.0048.

Author Contributions

Conceptualization, S.C.H., J.S.Y., P.H.Y., K.S.J.; data curation, all authors; formal analysis, S.C.H., J.S.Y., H.H.; funding acquisition, S.C.H.; investigation, J.S.Y.; methodology, S.C.H., J.S.Y., S.W.H., K.S.J.; project administration, S.C.H., J.S.Y.; resources, J.S.Y.; supervision, S.C.H.; software, S.W.H.; validation, S.C.H.; writing—original draft, J.S.Y., S.C.H.; and writing—review & editing, J.S.Y., S.C.H.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Funding

This work was supported by the National Research Foundation of Korea (NRF-2021R1C1C1014413).

REFERENCES

 Jack CR Jr. Alzheimer disease: new concepts on its neurobiology and the clinical role imaging will play. Radiology 2012;263:344-361

- Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement* 2018;14:535-562
- Park M, Moon WJ. Structural MR imaging in the diagnosis of Alzheimer's disease and other neurodegenerative dementia: current imaging approach and future perspectives. *Korean J Radiol* 2016;17:827-845
- 4. Lee JY, Park JE, Chung MS, Oh SW, Moon WJ. Expert opinions and recommendations for the clinical use of quantitative analysis software for MRI-based brain volumetry. *J Korean Soc Radiol* 2021;82:1124-1139
- 5. Brewer JB, Magda S, Airriess C, Smith ME. Fully-automated quantification of regional brain volumes for improved detection of focal atrophy in Alzheimer disease. *AJNR Am J Neuroradiol* 2009;30:578-580
- 6. Tanpitukpongse TP, Mazurowski MA, Ikhena J, Petrella JR; Alzheimer's Disease Neuroimaging Initiative. Predictive utility of marketed volumetric software tools in subjects at risk for Alzheimer disease: do regions outside the hippocampus matter? AJNR Am J Neuroradiol 2017;38:546-552
- 7. Wang C, Beadnall HN, Hatton SN, Bader G, Tomic D, Silva DG, et al. Automated brain volumetrics in multiple sclerosis: a step closer to clinical application. *J Neurol Neurosurg Psychiatry* 2016;87:754-757
- Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P, et al. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. J Neurol Neurosurg Psychiatry 1992;55:967-972
- Park HY, Park CR, Suh CH, Shim WH, Kim SJ. Diagnostic performance of the medial temporal lobe atrophy scale in patients with Alzheimer's disease: a systematic review and meta-analysis. *Eur Radiol* 2021;31:9060-9072
- Cavallin L, Løken K, Engedal K, Oksengård AR, Wahlund LO, Bronge L, et al. Overtime reliability of medial temporal lobe atrophy rating in a clinical setting. *Acta Radiol* 2012;53:318-323
- Park HY, Suh CH, Heo H, Shim WH, Kim SJ. Diagnostic performance of hippocampal volumetry in Alzheimer's disease or mild cognitive impairment: a meta-analysis. *Eur Radiol* 2022 May 4 [Epub]. https://doi.org/10.1007/ s00330-022-08838-9
- Enkirch SJ, Traschütz A, Müller A, Widmann CN, Gielen GH, Heneka MT, et al. The ERICA score: an MR imaging-based visual scoring system for the assessment of entorhinal cortex atrophy in Alzheimer disease. *Radiology* 2018;288:226-333
- Ochs AL, Ross DE, Zannoni MD, Abildskov TJ, Bigler ED; Alzheimer's Disease Neuroimaging Initiative. Comparison of automated brain volume measures obtained with NeuroQuant[®] and FreeSurfer. *J Neuroimaging* 2015;25:721-727
- 14. Persson K, Bohbot VD, Bogdanovic N, Selbaek G, Braekhus A, Engedal K. Finding of increased caudate nucleus in patients with Alzheimer's disease. *Acta Neurol Scand* 2018;137:224-232
- Lee JY, Oh SW, Chung MS, Park JE, Moon Y, Jeon HJ, et al. Clinically available software for automatic brain volumetry: comparisons of volume measurements and validation of intermethod reliability. *Korean J Radi*ol 2021;22:405-414
- 16. Kim M, Kim SJ, Park JE, Yun J, Shim WH, Oh JS, et al. Combination of automated brain volumetry on MRI and quantitative tau deposition on THK-5351 PET to support diagnosis of Alzheimer's disease. *Sci Rep* 2021; 11:10343
- Min J, Moon WJ, Jeon JY, Choi JW, Moon YS, Han SH. Diagnostic efficacy of structural MRI in patients with mild-to-moderate Alzheimer disease: automated volumetric assessment versus visual assessment. AJR Am J Roentgenol 2017;208:617-623
- 18. Noor MBT, Zenia NZ, Kaiser MS, Mamun SA, Mahmud M. Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. *Brain Inform* 2020;7:11
- 19. Korea Health Industry Development Institute. A study on the investigation of new medical device classification and management system in major countries. Cheongju: KHIDI 2021
- 20. Suh CH, Shim WH, Kim SJ, Roh JH, Lee JH, Kim MJ, et al. Development and validation of a deep learningbased automatic brain segmentation and classification algorithm for Alzheimer disease using 3D T1-weighted volumetric images. AJNRAm J Neuroradiol 2020;41:2227-2234
- **21.** Bae JB, Lee S, Jung W, Park S, Kim W, Oh H, et al. Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. *Sci Rep* 2020;10:22252
- 22. Lee M, Kim J, Kim R EY, Kim HG, Oh SW, Lee MK, et al. Split-attention U-Net: a fully convolutional network for robust multi-label segmentation from brain MRI. *Brain Sci* 2020;10:974



- 23. Kim REY, Lee M, Kang DW, Wang SM, Kim NY, Lee MK, et al. Deep learning-based segmentation to establish East Asian normative volumes using multisite structural MRI. *Diagnostics (Basel)* 2020;11:13
- 24. Song M, Jung H, Lee S, Kim D, Ahn M. Diagnostic classification and biomarker identification of Alzheimer's disease with random forest algorithm. *Brain Sci* 2021;11:453
- 25. Kim YJ, Han JW, So YS, Seo JY, Kim KY, Kim KW. Prevalence and trends of dementia in Korea: a systematic review and meta-analysis. *J Korean Med Sci* 2014;29:903-912
- 26. Erkinjuntti T. Clinical criteria for vascular dementia: the NINDS-AIREN criteria. Dementia 1994;5:189-192
- 27. Fazekas F, Kleinert R, Offenbacher H, Schmidt R, Kleinert G, Payer F, et al. Pathologic correlates of incidental MRI white matter signal hyperintensities. *Neurology* 1993;43:1683-1689
- 28. Noh Y, Lee Y, Seo SW, Jeong JH, Choi SH, Back JH, et al. A new classification system for ischemia using a combination of deep and periventricular white matter hyperintensities. J Stroke Cerebrovasc Dis 2014;23: 636-642
- 29. Mitchell T, Lehéricy S, Chiu SY, Strafella AP, Stoessl AJ, Vaillancourt DE. Emerging neuroimaging biomarkers across disease stage in Parkinson disease: a review. *JAMA Neurol* 2021;78:1262-1272
- 30. Pyatigorskaya N, Magnin B, Mongin M, Yahia-Cherif L, Valabregue R, Arnaldi D, et al. Comparative study of MRI biomarkers in the substantia nigra to discriminate idiopathic Parkinson disease. AJNR Am J Neuroradiol 2018;39:1460-1467
- **31.** Frosini D, Cosottini M, Volterrani D, Ceravolo R. Neuroimaging in Parkinson's disease: focus on substantia nigra and nigro-striatal projection. *Curr Opin Neurol* 2017;30:416-426
- **32.** Kim EY, Sung YH, Lee J. Nigrosome 1 imaging: technical considerations and clinical applications. *Br J Radiol* 2019;92:20180842
- 33. Kim PH, Lee DH, Suh CH, Kim M, Shim WH, Kim SJ. Diagnostic performance of loss of nigral hyperintensity on susceptibility-weighted imaging in parkinsonism: an updated meta-analysis. *Eur Radiol* 2021;31:6342-6352
- 34. Nam Y, Gho SM, Kim DH, Kim EY, Lee J. Imaging of nigrosome 1 in substantia nigra at 3T using multiecho susceptibility map-weighted imaging (SMWI). *J Magn Reson Imaging* 2017;46:528-536
- **35.** Shin DH, Heo H, Song S, Shin NY, Nam Y, Yoo SW, et al. Automated assessment of the substantia nigra on susceptibility map-weighted imaging using deep convolutional neural networks for diagnosis of Idiopathic Parkinson's disease. *Parkinsonism Relat Disord* 2021;85:84-90
- Hughes AJ, Daniel SE, Ben-Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain* 2002;125(Pt 4):861-870
- Höglinger GU, Respondek G, Stamelou M, Kurz C, Josephs KA, Lang AE, et al. Clinical diagnosis of progressive supranuclear palsy: the movement disorder society criteria. *Mov Disord* 2017;32:853-864
- 38. Gilman S, Wenning GK, Low PA, Brooks DJ, Mathias CJ, Trojanowski JQ, et al. Second consensus statement on the diagnosis of multiple system atrophy. *Neurology* 2008;71:670-676
- Quattrone A, Nicoletti G, Messina D, Fera F, Condino F, Pugliese P, et al. MR imaging index for differentiation of progressive supranuclear palsy from Parkinson disease and the Parkinson variant of multiple system atrophy. *Radiology* 2008;246:214-221
- 40. Chougar L, Faouzi J, Pyatigorskaya N, Yahia-Cherif L, Gaurav R, Biondetti E, et al. Automated categorization of parkinsonian syndromes using magnetic resonance imaging in a clinical setting. *Mov Disord* 2021;36:460-470
- **41.** Martín-Láez R, Caballero-Arzapalo H, López-Menéndez LÁ, Arango-Lasprilla JC, Vázquez-Barquero A. Epidemiology of idiopathic normal pressure hydrocephalus: a systematic review of the literature. *World Neurosurg* 2015;84:2002-2009
- **42.** Andersson J, Rosell M, Kockum K, Lilja-Lund O, Söderström L, Laurell K. Prevalence of idiopathic normal pressure hydrocephalus: a prospective, population-based study. *PLoS One* 2019;14:e0217705
- Relkin N, Marmarou A, Klinge P, Bergsneider M, Black PM. Diagnosing idiopathic normal-pressure hydrocephalus. *Neurosurgery* 2005;57(3 Suppl):S4-S16
- 44. Williams MA, Malm J. Diagnosis and treatment of idiopathic normal pressure hydrocephalus. *Continuum (Minneap Minn)* 2016;22:579-599
- **45.** Mori E, Ishikawa M, Kato T, Kazui H, Miyake H, Miyajima M, et al. Guidelines for management of idiopathic normal pressure hydrocephalus: second edition. *Neurol Med Chir (Tokyo)* 2012;52:775-809
- 46. Park HY, Kim M, Suh CH, Lee DH, Shim WH, Kim SJ. Diagnostic performance and interobserver agreement of the callosal angle and Evans' index in idiopathic normal pressure hydrocephalus: a systematic review and meta-analysis. *Eur Radiol* 2021;31:5300-5311

- **47.** Park HY, Park CR, Suh CH, Kim MJ, Shim WH, Kim SJ. Prognostic utility of disproportionately enlarged subarachnoid space hydrocephalus in idiopathic normal pressure hydrocephalus treated with ventriculoperitoneal shunt surgery: a systematic review and meta-analysis. *AJNR Am J Neuroradiol* 2021;42:1429-1436
- 48. Takagi K, Watahiki R, Machida T, Onouchi K, Kato K, Oshima M. Reliability and interobserver variability of evans' index and disproportionately enlarged subarachnoid space hydrocephalus as diagnostic criteria for idiopathic normal pressure hydrocephalus. *Asian J Neurosurg* 2020;15:107-112
- 49. Duan W, Zhang J, Zhang L, Lin Z, Chen Y, Hao X, et al. Evaluation of an artificial intelligent hydrocephalus diagnosis model based on transfer learning. *Medicine (Baltimore)* 2020;99:e21229
- 50. Gunter NB, Schwarz CG, Graff-Radford J, Gunter JL, Jones DT, Graff-Radford NR, et al. Automated detection of imaging features of disproportionately enlarged subarachnoid space hydrocephalus using machine learning methods. *Neuroimage Clin* 2019;21:101605
- 51. Kim DW, Jang HY, Kim KW, Shin Y, Park SH. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers. Korean J Radiol 2019;20:405-410
- Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology* 2018;286:800-809
- 53. Lee J, Lee JY, Oh SW, Chung MS, Park JE, Moon Y, et al. Evaluation of reproducibility of brain volumetry between commercial software, inbrain and established research purpose method, FreeSurfer. J Clin Neurol 2021;17:307-316

퇴행성 뇌질환에서 뇌 자기공명영상 기반 인공지능 소프트웨어 활용의 현재

정소영¹·서종현^{2*}·박호영²·허 훤²·심우현²·김상준²

현대사회가 점차 고령화 사회가 됨에 따라 퇴행성 뇌질환의 발병률이 증가하고 있으며, 이러 한 퇴행성 뇌질환에 관한 많은 연구들이 이루어지고 있다. 퇴행성 뇌질환의 진단에서 영상분 석은 영상표지자로서 중요한 역할을 하고 있다. 영상분석에서 객관적이고 일관성 있는 평가 는 퇴행성 뇌질환의 조기 진단 및 정확한 진단에 중요하다. 이에 다양한 퇴행성 뇌질환과 관 련한 영상연구에 자기공명영상(이하 MRI)을 이용한 인공지능이 조기 진단과 최적의 치료 방 향 계획 및 결정에 도움이 될 가능성을 보여주었다. 특히 MRI 기반의 뇌용적 측정과 분획화 및 특성을 포착하는 인공지능 소프트웨어들이 개발되고 연구되기 시작했다. 본 고찰에서는 우리나라에서 퇴행성 뇌질환과 관련하여 사용되고 있는 인공지능 소프트웨어의 현재 상황과 향후 인공지능 소프트웨어의 퇴행성 뇌질환 연구에의 활용, 그리고 인공지능 소프트웨어의 한계에 대해서 다루고자 한다.

¹한양대학교 의과대학 한양대학교병원 영상의학과, ²울산대학교 의과대학 서울아산병원 영상의학과