








# Expert Opinions and Recommendations for the Clinical Use of Quantitative Analysis Software for MRI-Based Brain Volumetry

뇌 자기공명영상 뇌용적 분석 소프트웨어의 임상적 적용에 대한 전문가 의견과 권고안

Ji Young Lee, MD<sup>1</sup> , Ji Eun Park, MD<sup>2</sup> , Mi Sun Chung, MD<sup>3</sup> ,  
Se Won Oh, MD<sup>4</sup> , Won-Jin Moon, MD<sup>5\*</sup> ; Aging and Neurodegeneration  
Imaging (ANDI) Study Group, Korean Society of Neuroradiology (KSNR)

<sup>1</sup>Department of Radiology, Hanyang University Medical Center, Hanyang University Medical College, Seoul, Korea

<sup>2</sup>Department of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

<sup>3</sup>Department of Radiology, Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul, Korea

<sup>4</sup>Department of Radiology, Eunpyeong St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea

<sup>5</sup>Department of Radiology, Konkuk University Medical Center, Konkuk University School of Medicine, Seoul, Korea

The objective assessment of atrophy and the measurement of brain volume is important in the early diagnosis of dementia and neurodegenerative diseases. Recently, several MR-based volumetry software have been developed. For their clinical application, several issues arise, including the standardization of image acquisition and their validation of software. Additionally, it is important to highlight the diagnostic performance of the volumetry software based on expert opinions. We instituted a task force within the Korean Society of Neuroradiology to develop guidelines for the clinical use of MR-based brain volumetry software. In this review, we introduce the commercially available software and compare their diagnostic performances. We suggest the need for a standard protocol for image acquisition, the validation of the software, and evaluations of the limitations of the software related to clinical practice. We present recommendations for the clinical applications of commercially available software for volumetry based on the expert opinions of the Korean Society of Neuroradiology.

**Index terms** Magnetic Resonance Imaging; Software; Diagnosis; Brain; Gray Matter

Received October 7, 2020  
Revised December 31, 2020  
Accepted January 21, 2021

**\*Corresponding author**

Won-Jin Moon, MD  
Department of Radiology,  
Konkuk University  
School of Medicine,  
Konkuk University Medical Center,  
120-1 Neungdong-ro,  
Gwangjin-gu, Seoul 05030, Korea.

Tel 82-2-2030-5544

Fax 82-2-2030-5499

E-mail mdmoonwj@naver.com

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**

Ji Young Lee 


[https://](https://orcid.org/0000-0003-1181-8070)

[orcid.org/0000-0003-1181-8070](https://orcid.org/0000-0003-1181-8070)

Ji Eun Park 

[https://](https://orcid.org/0000-0002-4419-4682)

[orcid.org/0000-0002-4419-4682](https://orcid.org/0000-0002-4419-4682)

Mi Sun Chung 


[https://](https://orcid.org/0000-0003-1141-9555)

[orcid.org/0000-0003-1141-9555](https://orcid.org/0000-0003-1141-9555)

Se Won Oh 

[https://](https://orcid.org/0000-0003-1336-4498)

[orcid.org/0000-0003-1336-4498](https://orcid.org/0000-0003-1336-4498)

Won-Jin Moon 

[https://](https://orcid.org/0000-0002-8925-7376)

[orcid.org/0000-0002-8925-7376](https://orcid.org/0000-0002-8925-7376)

## 서론

퇴행성 신경계 질환 및 치매 진단에 뇌 자기공명(이하 MR) 영상의 구조적 영상뿐 아니라, MR 영상에 기반한 뇌용적 측정(brain volumetry)의 정량화 분석(quantitative analysis)이 유용하게 쓰이고 있다. 알츠하이머 치매를 비롯한 퇴행성 신경 질환에서 정량화 분석을 이용한 연구들은 특정 부위의 뚜렷한 뇌 위축이 정상인과 치매 환자를 구별하고 병변의 진행정도를 파악할 수 있는 영상표지자임을 보고하였다(1-3). 그러나, 뇌용적 측정 영상분석 기술발달에도 불구하고, 뇌용적 측정(volumetry)은 분석에 시간이 많이 요구되어 주로 연구용으로 이용되어 왔다(2, 3). 실제 뇌 위축 평가는 영상의학과 전문의의 시각적 평가에 주로 의존하고 있고(4), 이러한 시각적 평가는 관찰자 내 일치도는 매우 높으나, 서로 다른 관찰자 간 일치도는 0.29-0.48로 낮고 변동폭이 크며(4), 또한 전문가와 비전문가 간의 차이도 매우 크다(5).

최근 소개되는 임상용 정량 분석 소프트웨어들은 뇌의 회색질 및 백질의 부피를 자동 정량 분석하여, 특정 뇌영역의 위축 정도를 평가하고 소프트웨어 내에 저장되어 있는 정상규준자료(normative data)를 이용하여, 정상규준 백분위수(normative percentile)와 같은 통계수치를 제공하고 있어 시각적 평가의 문제점을 해결할 수 있다. 현재 관련 시스템 및 소프트웨어들이 경쟁적으로 개발되고 있으며, 2018년 8월까지 적어도 3가지 이상의 임상용 소프트웨어가 미국 식품의약안 신청에서 임상용으로 승인을 받아, 정식으로 보험청구 가능한 항목으로 인정을 받고 있다(6-8).

기존의 연구용 소프트웨어와 비교하여 임상용 소프트웨어들은 쉬운 사용자 인터페이스를 공급하고 별다른 전문지식이나 복잡한 분석 과정 없이 받아볼 수 있는 직관적인 결과보고서를 제공하여 환자나 의사 모두에게 매우 매력적이다. 그러나 뇌 MR 영상은, 본질적으로 절대수치가 아닌 상대수치의 조합이기 때문에, 표준화가 되지 않은 원본 영상데이터를 투입하는 경우 발생할 수 있는 문제점에 대한 인식과 책임에 대해 전문가들의 논의가 이루어지지 않았다. 따라서 전문가의 검증과 합의를 거친 뚜렷한 권고안이 없는 상태에서 무분별하게 이용하는 경우, 임상용 소프트웨어 간의 분석 결과값 차이에 따른 뇌 위축의 위양성, 위음성 등의 부작용이 발생할 수 있다. 또한, 주자기장강도(magnetic field strength) 차이, MR 하드웨어의 차이, MR 영상 획득인자 값의 차이로 발생하는 임상용 소프트웨어 결과 차이에 대한 고찰과 함께 MR 영상 획득의 표준화가 필요하다고 판단된다. 향후 인공지능과 빅데이터 관련 기술의 급속한 발달은 이러한 소프트웨어의 개발과 임상 적용을 가속화할 것이기 때문에, 전문가에 의한 적절한 평가와 전문가 의견 마련이 시급하다. 이에 대해 대한신경두경부영상의학회(Korean Society of Neuroradiology; 이하 KSNR)는 뇌 MR 영상의 임상용 정량화 분석 소프트웨어의 사용에 있어서 현재까지 발표된 소프트웨어들의 진단능력에 대한 근거수준을 정리하고, 영상 획득 표준화, 임상용 소프트웨어 진단능에 대한 평가 및 실제 임상 적용에 대한 전문가 권고안을 제시하기 위하여 산하 퇴행성신경질환연구회를 중심으로 권고안 개발 전문위원회를 구성하여 지침 작업을 시작하였다. 현재까지의 연구 자료들의 방법론적 이질성과 충분하지 않은 근거 자료들로 인해, 본 종설에서는 향후 보다 정확하고 엄밀한 임상 지침을 개발하기 전 단계로서 전문가견해를 요약한 결과를 소개하기로 하고, MR 영상을 이용한 뇌용적 측정 소프트웨어 이용에 대하여 기존의 표준품(reference standard)과의 비교 및 신

뢰도와 재현성 연구 결과를 정리하고, 임상 적용에 대한 제한점 및 대처 방법에 대한 전문가 권고안을 소개하고자 한다.

## 연구 방법

### 평가위원 선임과 과정

KSNR 회원이며 신경영상진단에 오랜 진료경험이 있으며 정량화 분석과 소프트웨어 적용에 대한 축적된 경험이 있는 전문가 혹은 수련병원에서 신경영상진단을 담당하고 있는 전문가 14명을 권고안 개발 전문위원회에 포함하였다. 영상의학과 전문가 구성은 대학병원당 1-2인을 원칙으로 하여 가능한 다수의 수련병원이 참여할 수 있도록 하였다. 또한, MRI 영상에 전문적인 지식이 있는 비의사 영상전문가(PhD) 1인을 포함하여 총 15명의 전문가가 전문위원회에 참여하였다.

핵심 내용에 대한 합의는 대면회의를 통하여 결정하였으며, 필요시 이메일로 의견을 교환하여 결정하였다. 지정된 담당자가 기본 고찰 및 안을 제시하고, 해당 주제에 관하여 전문 위원들의 토론 후 결정하였다. 또한, 기존의 근거가 부족한 임상용 소프트웨어 간의 타당성 및 재현성 검증을 위하여 대표적인 임상 소프트웨어 간의 비교 연구를 전문위원회를 중심으로 국내 영상자료를 이용하여 진행하였다.

### 문헌선정 및 집필 과정

본 연구를 위하여 2009년 이후부터 2017년 5월까지 발표된 뇌 자기공명영상의 정량화 분석 및 임상용 소프트웨어 관련 연구 문헌과 진료 관련 지침들을 사전 분석 고찰하였다. 한국보건의료연구구원(National Evidence-based Healthcare Collaborating Agency)의 도움을 받아 관련 문헌을 검색하여, 1차로 PUBMED에서 총 186건, EMBASE에서 280건, COCHRANE에서 28건의 문헌이 검색되었으며, 중복 문헌을 제거 후 총 434개 문헌을 선정하였다. 1차 선정된 434개 문헌을 각 6그룹으로 나누어 각 그룹당 2명의 검토자가 초록을 검토한 다음, 292건의 문헌을 2차 선정하여 전체 그룹이 재검토하고, 다시 3차 선정된 문헌 115건에 대해 초록 및 본문 검토하여 51개 문헌을 선정하였다. 또, 2017년 6월부터, 2018년 6월까지 추가로 본 주제와 관련된 문헌 13건을 추가하여 최종적으로 64개의 문헌이 포함되었다(Supplementary Table 1 in the online-only Data Supplement).

이를 바탕으로 2018년 11월 20일 권고안 초안을 KSNR 집담회에서 중간보고하였으며, 2019년 전문위원회를 중심으로 진행한 연구 결과를 추가하고, 2018년 7월부터 2020년 6월까지의 관련 문헌 5건을 추가하여, 권고안을 최종적으로 수정 정리하였다(9-13).

### 임상용 뇌용적 측정 정량화 분석 소프트웨어 현황

뇌용적 정량적 분석은 정성적 분석과 비교해 뇌 위축 정도에 대해 객관적인 수치를 제공해주고, 대뇌 백질, 회백질, 전체 뇌용적에 대한 자세한 수치를 제공해주는 장점이 있다. 뇌용적의 정량적 분석이 가능한 프로그램으로 문헌상 가장 널리 인용되는 것은 FMRIB Software Library (이하 FSL) (14), FreeSurfer (15), Statistical Parametric Mapping (이하 SPM) (16) 등이 있는데 이는

주로 연구용 소프트웨어로서, 이용자가 임상적 목적으로 사용하기에는 분석 시간이 오래 걸리고, 임상가들이 사용하기 어렵고, 환자 진단을 위한 목적으로 사용할 수 있는 허가를 획득하지 않았다는 문제점이 있다(17).

현재 임상용으로 허가받은 뇌용적 분석 소프트웨어의 목록은 Table 1과 같다. 그중 가장 널리 쓰이고 문헌에서 확인할 수 있는 소프트웨어는 다음 4가지이다.

### NeuroQuant<sup>®</sup> (CorTechs Labs, San Diego, CA, USA)

2009년 개발된 이후로, 임상용 소프트웨어 중 가장 많이 사용되고 있고, 여러 연구들을 통해 비교적 많은 검증이 이루어진 소프트웨어이다(6, 12, 18-24). 웹 기반 정량적 분석 소프트웨어로서, 약 5-10분 이내로 결과를 얻을 수 있다(9). 이는 각각의 뇌이랑(gyrus)에 대한 회백질과 백질의 용적을 제시하는 대신에 전체 백질과 전체 회백질의 용적을 제시해주는 방식과 관련이 있다. PACS에서 결과보고서를 바로 제공받을 수 있고, 환자의 데이터를 소프트웨어에 내장된 정상인 데이터와 비교한 정상규준 백분위수를 보여줄 수 있다. NeuroQuant<sup>®</sup>는 전체 뇌용적과 11 부위(편도체, 미상핵, 소뇌, 대뇌피질, 전뇌실질, 해마, 하부측뇌실, 측뇌실, 창백핵, 조가비핵, 시상)의 용적을 양측에 대해 각각 제시해준다. NeuroQuant<sup>®</sup>는 FreeSurfer에 기반하여 개발하였으나, 구획 나누기에는 서로 다른 해부학 아틀라스를 사용하고, 독립적이고 개별적인 방식을 이용하여 결과를 얻어낸다(18). NeuroQuant<sup>®</sup>는 뇌용적 분석의 정확성을 위하여, 소프트웨어에서 추천하는 MR 영상 획득 인자를 사용하기를 권고한다(Supplementary Fig. 1 in the online-only Data Supplement).

### Neuroreader<sup>®</sup> (Brainreader Aps, Horsens, Denmark)

프로세싱 시간이 10분 정도 소요되며, NeuroQuant<sup>®</sup> 보다 더 많은 64개의 뇌구조물에 대한 용적을 제시해준다. 비선형적 정합(nonlinear registration)을 통해, 여러 아틀라스에 기반한 뇌분획화 방법을 이용한다(7, 25). NeuroQuant<sup>®</sup>와 마찬가지로, PACS에서 결과보고서를 바로 제공받을 수 있고, 환자의 데이터를 소프트웨어 내의 정상인 데이터와 비교한 정상규준 백분위수를 보여줄 수 있다(Supplementary Fig. 2 in the online-only Data Supplement).

### Icobrain (Icometrix, Leuven, Belgium)

최근까지 문헌들에서 MSmetrix로 소개되었으며, 다른 소프트웨어와 달리, 원래 다발성경화증 환자를 위해 고안된 소프트웨어로서 전체 뇌용적, 회백질, 백질, 뇌척수액, 피질, 백질 고신호강도 영역의 용적에 대한 결과를 제공해주며, 반드시 T1강조영상과 fluid attenuated inversion recovery (이하 FLAIR) 영상 둘 다 있어야 결과가 나오며, 약 30분의 시간이 소요된다. 환자의 추적검사 영상을 이용하여, 종적인 분석도 가능하다(11, 26-29). 다발성경화증의 병변 변화를 추적하기 위해 개발되어, 다양한 영역의 백질 병변을 구분하는데 뛰어나고, 치매 관련 영역에서는 상대적으로 적은 수의 뇌 영역만을 구분하여 용적을 구해낸다(Supplementary Fig. 3 in the online-only Data Supplement).

Table 1. Commercially Available Brain Volumetry Software

Software	FDA Approved	MFDS Approved	EU CE Marked	URL	Country	Company	Character
NeuroQuant®	FDA approved	NA	CE marked	<a href="https://cortechs.ai/">https://cortechs.ai/</a>	USA	Cortechs Lab	NeuroQuant, LesionQuant*
Neuroreader®	FDA approved	NA	CE marked	<a href="https://brainreader.net/">https://brainreader.net/</a>	USA	Brainreader	
Icobrain	FDA approved	NA	CE marked	<a href="https://icomatrix.com/">https://icomatrix.com/</a>	EU	Icomatrix	Icobrain MS, Icobrain DM, Icobrain TBI, Icobrain ep
Quantib® Brain	FDA approved	NA	CE marked	<a href="https://www.quantib.com/">https://www.quantib.com/</a>	Netherlands	Quantib	
syMRI® Neuro	FDA approved	NA	CE marked	<a href="https://syntheticmr.com/">https://syntheticmr.com/</a>	Sweden		Multicontrast† Volumetry (GM, WM, CSF volume)
InBRAIN®	NA	MFDS approved	NA	<a href="https://www.inbrain.co.kr/">https://www.inbrain.co.kr/</a>	South Korea	Midas IT	
Neurol	NA	MFDS approved	NA	<a href="http://www.infomeditech.com/">http://www.infomeditech.com/</a> <a href="https://neurozen.ai/">https://neurozen.ai/</a>	South Korea	Infomeditech	
mdbrain	NA	NA	CE marked	<a href="https://www.qmenta.com">https://www.qmenta.com</a>	EU	Qmenta Inc	
DeepBrain®	NA	MFDS approved	CE marked	<a href="https://www.vuno.co/">https://www.vuno.co/</a>	South Korea	Vuno	
Atroscan	NA	MFDS approved	NA	<a href="http://jlkgroup.com">http://jlkgroup.com</a>	South Korea	JLK	

\*LesionQuant: measurement of WM lesion volume.

†Multicontrast: acquisition of multiple pre-defined contrast weighted images such as T1W, T2W and T2W FLAIR, as well as double inversion recovery and phase sensitive inversion recovery using synthetic MRI.

CSF = cerebrospinal fluid, EU CE Marked = conforme Europeenne marked, FDA = Food and Drug Administration, FLAIR = fluid attenuated inversion recovery, GM = grey matter, MFDS = Ministry of Food and Drug Safety, NA = not applicable, T1W = T1-weighted, T2W = T2-weighted, WM = white matter

### Inbrain® (MIDAS IT, Seongnam, Korea)

2017년 식품의약품안전처(Ministry of Food and Drug Safety) 승인받은 뇌용적 분석 소프트웨어로서, 기본적으로 FreeSurfer 플랫폼 기반으로 하여 FreeSurfer와 동일한 해부학적 아틀라스를 이용하여 분획화를 시행한다(30). 뇌 분획화 및 백질 분획화 단계에서 부분적으로 딥러닝 기법이 적용되고, 피질 두께 측정까지 시행하므로, 결과를 얻는데 약 4시간 정도 소요된다. 결과적으로, 대뇌부피, 회백질, 백질, 소뇌, 외측내실, 전두엽, 두정엽, 측두엽, 후두엽, 미상핵, 조가비핵, 편도체, 해마, 중격핵(nucleus accumbens), 창백핵, 시상 대뇌피질두께, 백질 고신호강도에 대해 정량적 분석 결과를 제시한다(Supplementary Fig. 4 in the online-only Data Supplement).

### 뇌용적 측정 정량화 분석 소프트웨어의 신뢰도와 재현성

자기공명영상 검사를 통한 정량 분석의 신뢰도는 크게 네 가지로 평가할 수 있다. 첫 번째는 한 환자가 짧은 일정 시간 간격을 두고 같은 기기에서 촬영했을 때 같은 정량 분석 결과가 얻어지는지 확인하는 반복검사의 신뢰도 혹은 반복성(test-retest reliability, 즉 test-retest repeatability)이다(11, 31). 두 번째는 같은 기기에서 다른 영상 기법으로 얻었을 때 같은 정량 분석 결과가 도출되는지를 확인하는 영상 획득 재현성(imaging acquisition reliability, sequence 간 reproducibility)이다(10, 11, 32). 세 번째는 같은 환자가 다른 자기공명영상 기기에서 촬영했을 때에도 같은 영상이 나오는 확인하는 기기 간 재현성(inter-scanner reproducibility)이다. 마지막으로 서로 다른 관찰자(reader), 소프트웨어를 통한 영상 후 처리 및 분석에서 같은 정량적 지표가 나오는지 확인하는 분석자 간 재현성(inter-reader or inter-method reproducibility)이다(10, 31). 앞의 세 가지는 영상 자체의 반복성과 재현성, 네 번째 항목은 영상 후처리 정량화 분석의 재현성 측면으로 볼 수 있다. 정량화 분석에서는 주로 영상 후처리 정량화 분석을 시행하는 임상용 소프트웨어 간 재현성이 강조되는데, 자기공명영상 자체의 반복성과 재현성도 중요하다.

### 연구용 정량 분석 소프트웨어의 재현성 연구 문헌고찰

Fellhauer 등(33)은 연구에서 디지털 팬텀, 34명의 알츠하이머 치매 환자, 60명의 경도인지 장애 환자, 32명의 정상인을 대상으로 한 연구에서 FreeSurfer, FSL, SPM을 이용한 회색질과 백색질의 정량적 용적 측정을 비교하였다. 디지털 팬텀연구에서, 영상의 잡음이 증가함에 따라, SPM, FSL, FreeSurfer 모두 회색질 용적은 과소 측정되고, 백색질 용적은 과대 측정되는 경향이 있었다. 영상이 비균질한 경우, SPM과 FreeSurfer는 비교적 영향을 받지 않았으나, FSL의 경우는 백색질 용적은 증가하고 회색질 용적은 감소하는 경향을 보였다. 환자 영상 결과에서 세 소프트웨어 모두 나이에 따른 뇌 위축에 의한 회색질과 백색질 감소를 뚜렷하게 보여주었다. 세 소프트웨어 중에서 FreeSurfer가 회색질 용적을 가장 크게 측정하였고 백색질 용적은 가장 작게 측정하였다. FSL은 가장 적은 회색질 용적을 보였다. SPM은 가장 큰 백색질 용적을 보였다. 영상의 품질이 나쁜 경우에는 SPM이 최상의 분획화 결과를 제공하였다.

Mulder 등(34)이 해마 용적의 측정을 비교한 연구에서는 1.5T에서 1년 간격으로 각각 2번씩 얻은 Alzheimer's disease neuroimaging initiative (이하 ADNI) 공공데이터의 T1 강조영상들을 이



용하여 FreeSurfer (5.1.0)와 FSL-FIRST (4.1.5), 그리고 수기분획법(manual segmentation)을 비교하였다. 분획의 실패는 FreeSurfer는 약 6.4%, 와 FSL-FIRST는 약 5.3%였다. 분획 실패의 원인은 전체 뇌용적이 감소한 경우, 뇌실의 증가 및 백질 내 이상 소견이 있는 등이었다. 실패한 증례를 제외하고, 1년간 용적 변화율(%)의 재현성을 구했을 때, 일치한계(limits of agreement)는 FreeSurfer  $\pm$  7.2%, 수기분획  $\pm$  9.7%와 FSL-FIRST  $\pm$  10.0%로 수기분획과 FSL-FIRST 간은 큰 차이가 없었으나, FreeSurfer가 의미 있게 재현성이 뛰어났다.

2016년 비슷한 해마 용적 측정의 서로 다른 소프트웨어 간 신뢰도 및 재현성 비교 연구에서는 1.5T ADNI 공공데이터를 이용하여 FreeSurfer (5.3.0), FSL-FIRST (5.0.4), AdaBoos라는 머신러닝기반 분석법, multiple-atlas propagation and segmentation with hippocampal boundary shift integral (이하 MAPS-HBSI)을 이용하여 비교하였다(35). 이 연구는 한 시점에 두 번 촬영한 (back-to-back) 영상을 이용하여 각각 두 영상에서 측정한 해마 용적 차이를 구하여, 처음 측정한 해마 용적으로 나누어 비율(%)을 구하였다. 그 결과 MAPS-HBSI가 1.3%로 해마 용적 차이 비율이 가장 낮았고, 따라서 위 방법 중에서 가장 재현성이 뛰어났다.

한편 Nestor 등(36)의 연구에서는 5개의 서로 다른 뇌 아틀라스를 이용한 자동 해마 용적 측정을 비교하였으며, 좀 더 후측 해부와 배측 뇌백질(caudal hippocampus와 alveus-fimbria)을 포함할수록 수기분획과의 일치도가 좋아짐을(Dice similarity coefficient = 0.87-0.89) 보여주었다. 따라서, 기준이 되는 수기분획과 비교할 때, 어떤 원칙으로 자동 분획하는지에 따라 정확도가 달라짐을 증명하였다(36).

### 임상용 정량적 분석 소프트웨어의 재현성 연구 문헌고찰

NeuroQuant<sup>®</sup>는 제일 먼저 개발된 임상용 소프트웨어이다. NeuroQuant<sup>®</sup>를 이용한 20명의 치매 환자와 20명의 정상인을 대상으로 한 연구에서 치매 환자의 해부학적 위축을 민감하게 찾아 내었고 NeuroQuant<sup>®</sup> 결과를 수기분획 결과와 비교했을 때, 해마 등 어려운 부위에서도 급내상관계수(intraclass correlation coefficient; 이하 ICC)가 0.93으로 우수한 일치 결과를 보여주었다(6). 148명의 경도뇌손상환자 및 손상후증후군을 대상으로 한 연구에서는 NeuroQuant<sup>®</sup>와 FreeSurfer를 비교하였으며, NeuroQuant<sup>®</sup>와 FreeSurfer 두 소프트웨어 간의 급내상관계수는 0.4에서 0.99를 보이며 넓은 범위의 일치도를 보였으나, 30개 뇌 영역 중 26개 영역에서 두 소프트웨어의 결과값의 평균치를 비교하였을 때, 통계적으로 유의하게 차이가 나서, 임상적인 이용에 주의를 기울여야 함을 환기시켰다(19). 다발성경화증에 임상용 영상소프트웨어인 NeuroQuant<sup>®</sup>와 MSmetrix (현 Icobrain), 연구용 소프트웨어인 SIENAX의 비교 분석에서 임상용 영상소프트웨어와 연구용 소프트웨어 간 용적 분석의 차이가 1.0%-5.5%로 발생하였음을 보고하였다(8).

Steenwijk 등(27)은 20개의 모의뇌(simulated brain)와 100명의 다발성경화증 환자, 20명의 정상인을 대상으로 임상용 소프트웨어 MSmetrix를 연구용 소프트웨어 FreeSurfer, SIENAX 및 SPM과 비교하였다. 모의뇌 자료에서 SIENAX가 기준 전체 뇌용적과 가장 적은 차이를 보였고, 그 다음은 MSmetrix, SPM, FreeSurfer 순이었다. FreeSurfer와 SPM은 뇌용적을 과소 측정하였고, SIENAX는 과대 측정하였다.

종단 데이터 비교는 MRI 하드웨어의 업그레이드에 민감하였다. MSmetrix를 다발성경화증의 뇌 위축과 회색질위축을 평가하기 위하여, 모의뇌데이터, 단면데이터와 종단데이터를 이용하여 ANTs-v1.9, CIVET-v2.1, FSL-SIENAX/SIENA-5.0.1, SPM12과 비교한 연구에서, 전체 뇌와 회색질 용적에는 모두 높은 정확도를 보였으며, MSmetrix가 그중 가장 낮은 정확도를 보였다. 전반적으로 모든 방법에 대하여 높은 반복성을 보였으나, 낮은 재현성을 보였다. 따라서 추적검사 시 MR 영상 기기의 변경은 가급적 피해야 하며, 좀 더 표준화된 MR 영상 획득이 필요하다고 제안했다.

## KSNR 전문위원회의 임상용 소프트웨어와 연구용 소프트웨어 비교 연구

기존 문헌에서는 모두 서양인의 영상데이터를 기반으로 하였고, 주로 이미 질환이 진행된 환자를 중심으로 연구가 진행되어, 경미한 임상 증상이 있는 환자들에 대한 연구는 아직도 부족한 실정이다. 또한, 영상 프로토콜 특히 절편 두께의 차이에 따른 용적 측정 결과의 변화에 대하여 명확히 알려진 바가 없고, 국내 개발 소프트웨어의 재현성 연구가 없다는 한계가 있었다. 국내에서 가장 많이 이용되는 연구용 소프트웨어와 임상용 소프트웨어 중, NeuroQuant<sup>®</sup>, 그리고 국내 식품의약품안전처의 인증받은 Inbrain<sup>®</sup>을 중심으로 비교 연구 결과를 소개한다.

### FreeSurfer와 NeuroQuant<sup>®</sup>

FreeSurfer는 임상적 사용보다는 연구 목적으로 사용되고 있는 소프트웨어로서, 여러 뇌용적 분석 소프트웨어들의 비교 분석 연구들에서 상대 기준으로 사용되고 있는 소프트웨어이다. 공개 소스로 이용이 가능하나, 시간이 많이 걸리고 임상가가 다루기 어렵다는 단점이 있다. 80명의 경도인지장애 환자를 대상으로, FreeSurfer와 NeuroQuant<sup>®</sup>를 비교하였을 때, 전체 뇌용적, 대뇌피질, 백질 부위에서 NeuroQuant<sup>®</sup>가 FreeSurfer 보다 용적을 좀 더 크게 측정하는 경향을 보였다. 대부분의 영역에서 두 방법 간의 신뢰도는 좋은 편이다( $0.72 < ICC < 0.96$ ,  $0.78 < \text{Pearson correlation coefficient} < 0.94$ ). 그러나 대뇌하 피질 구조(조가비핵, 창백핵, 시상)의 부위는 두 방법 간의 신뢰도가 매우 나쁜 결과를 보여줬는데(37), 이는 기존에 발표되었던 Ochs 등(18)의 연구와 일치한다.

### FreeSurfer와 Inbrain<sup>®</sup>

130명의 정상인 및 경도인지장애 환자들을 대상으로 후향적 연구(38)를 시행하였을 때, 두 소프트웨어의 신뢰도는 매우 좋은 편이었다( $0.93 < ICC < 0.99$ ). 변동계수, coefficient of covariance의 통계적 방법에 따르면, 측좌핵(nucleus accumbens)과 같은 작은 해부학 구조물의 경우에 변동계수가 큰 것을 보여주었다. 또, 43명에서 1 mm 절편 두께로 촬영한 3D T1 강조영상으로부터 얻은 두 소프트웨어의 재현성과 42명에서 1.2 mm 절편 두께로 얻은 3D T1 강조영상의 소프트웨어 재현성을 비교하였을 때, 부위별로 다른 결과를 보여주었다. 측좌핵의 경우, 1 mm로 영상을 획득했을 때, FreeSurfer와 Inbrain<sup>®</sup> 두 소프트웨어 간의 재현성이 좋았다. 그러나, 피질두께의 경우에는 1.2 mm로 영상을 획득했을 때, FreeSurfer와 Inbrain<sup>®</sup> 두 소프트웨어 간의 재현성이



더 좋았다. 따라서, 향후 다기관 연구를 통해서 소프트웨어의 재현성에 영향을 주는 MR 영상 획득 인자에 관한 연구가 필요할 것으로 생각된다.

### NeuroQuant<sup>®</sup>와 Inbrain<sup>®</sup>

172명의 정상인, 경도인지장애 환자, 치매 환자를 대상으로, 두 소프트웨어로 얻은 용적을 비교하였을 때 대부분의 대뇌 부위에서 두 방법 간에 통계학적으로 의미 있는 용적 차이가 있었다. 두 소프트웨어의 차이를 표준화된 평균 차이로 보여주는 효과 크기(effect size)의 통계 방법에 따라 비교하였을 때, 두 소프트웨어의 차이는 큰 것으로 나타났다( $0.05 < \text{effect size} < 6.15$ ). 두 방법 간의 신뢰도는 대부분 대뇌 영역에서 좋은 편이었다( $0.83 < \text{ICC} < 0.98$ ,  $0.72 < \text{Pearson correlation coefficient} < 0.96$ ). 그러나 담창구(globus pallidus) 부위는 두 소프트웨어 간에 차이가 매우 컸고 매우 낮은 신뢰도를 보여주었다(39).

결론적으로, FreeSurfer, Inbrain<sup>®</sup>, NeuroQuant<sup>®</sup> 소프트웨어를 비교했을 때, 소프트웨어 간 신뢰도는 좋은 편이나, 그 측정 결과는 소프트웨어마다 의미 있는 차이를 보였다. 특히, 피질하 구조들에서는 그 측정 결과값이 소프트웨어마다 다를 수 있기에 임상에서 환자에게 적용하기 위해서는 검증 및 표준화가 필요하며, 이에 대해 관련 전문가에게 보고해야 할 것으로 생각된다.

### 적용 질환에서의 진단 정확도 연구

임상용 소프트웨어를 이용한 다양한 신경 질환의 진단 정확도에 대한 연구는 Table 2에 정리하였다. 20명의 치매 환자와 20명의 정상인을 대상으로 한 초기 연구에서 치매 환자의 해부학적 위축을 민감하게 찾아내었다(6). 치매 환자와 정상인에 임상용 소프트웨어를 적용한 국내 연구에서, 치매 환자의 특정 부위의 해부학적 위축을 전문가 판단과 비교하여 대등한 진단 정확도를 보였지만, 정상인에서 영상인공물로 인한 위양성이 있었다(40). 다른 문헌에서는 해마 용적의 전자동 정량 분석 소프트웨어의 알츠하이머 치매로의 진행 예측도가 0.68-0.69로 보고하였다(7).

전간증의 원인인 내측두경화증(mesial temporal sclerosis)의 진단은 신경영상전문의와 전자동 정량 분석의 진단 정확도는 각각 72.6%와 79.4%로 보고하였다(41). 최근 연구에서 해마경화의 진단에 신경영상전문의의 시각적 진단과 비슷한 진단 특이도(90.4% vs. 91.6%,  $p = 0.99$ )를 보였지만, 낮은 민감도(69.0% vs. 93.0%,  $p < 0.001$ )를 나타냈다(9).

### 정량적 분석의 오류 및 제한점

정량적 분석의 오류는 물리적 제한, 한국인 대상 대용량 정상데이터 부재, 마지막으로 병태생리적 제한에서 기인할 수 있다.

물리적 제한점은 대표적으로 피할 수 없는 인공물에 의한 정량 측정의 오류를 들 수 있다. Min 등(40)은 해마를 가로지르는 둘러겹침 인공물(aliasing artifact)로 인해 해마 용적이 과소 측정되어 뇌 위축 위양성으로 판정이 날 수 있음을 보고하였다.

Table 2. Research on the Diagnostic Performances of Commercially Available Volumetry Software

References	Structure	Purpose	Method	Subjects	Accuracy*	Sensitivity*	Specificity*	AUC
Azab et al. (2015) (41)	Hippocampus	Medial temporal sclerosis	NeuroQuant	Hippocampal sclerosis	73.8	NA	NA	NA
Min et al. (2017) (40)	Medial temporal lobe	Comparison between NeuroQuant and visual assessment	NeuroQuant	30 AD, 25 HC	NA	63.3	100	0.8
Tanpitukpongse et al. (2017) (7)	Whole brain	Assessment of prognostic efficacy of regional volume	NeuroQuant and Neuroreader	85 AD converter, 107 MCI	NA	NA	NA	0.6
Niemantsverdriet et al. (2018) (29)	Whole brain, GM, WM, CSF, cortical GM volumes, WM hyperintensities	Prediction of clinical progression	MSmetrix	313 AD, 379 MCI, 102 cognitive decline, 93 HC	NA	85.6	78.8	0.8
Perrson et al. (2018) (21)	Hippocampus	Comparison between NeuroQuant and visual assessment	NeuroQuant	31 AD, 25 HC	82	74.0	92.0	0.8
Louis et al. (2020) (9)	Hippocampus	Medial temporal sclerosis	NeuroQuant	144 temporal lobe epilepsy	NA	69.0	90.4	NA

\*Means percentile.  
 AUC = area under the curve, AD = Alzheimer's disease, CSF = cerebrospinal fluid, GM = grey matter, HC = healthy controls, MCI = mild cognitive impairment, NA = not applicable, WM = white matter

MR 하드웨어의 차이, 특히 주자기장강도의 차이는 매우 큰 영향을 주어, 종단연구에서는 반드시 같은 자기장강도의 MR 영상에서 획득한 데이터만을 이용하는 것을 권고한다. 반면, MR 시스템 업그레이드에 대해서는 알려진 바가 적다. Potvin 등(42)은 3명의 정상인을 대상으로 한 예비연구에서, Magnetom Trio를 Prisma로 업그레이드한 후 뇌피질의 대조도 대 잡음비(contrast-to-noise ratio; 이하 CNR)이 30% 증가한 반대급부로 뇌용적과 뇌피질 두께가 (특히 전두엽에서) 더 크게 측정됨을 보고하였다. FLAIR 영상에서는 CNR 차이가 관찰되지 않았지만, 뇌백질 고신호강도 병변(white matter hyperintensity)의 용적이 68% 더 적게 측정됨을 보고하였다.

2019년 Lee 등(43)은 237명의 1.5T로 시행한 ADNI 공공데이터를 이용하여 SIENAX로 종단 전 뇌 용적 변화를 측정하는 데 있어 MR 기기의 영향을 비교하였다. 이 연구는 GE, Philips, Siemens 3개의 MR 제조사 간의 차이를 보여주었다. GE Signa에서 Philips Intera는 0.99%, Philips Intera에서 Siemens Avanto는 -1.81% 변화가 나타남을 보여주었으며, 이는 경도인지장애에서 Alzheimer's disease로의 전환 시 1년간의 용적 변화에 해당하는 정도이다.

Magnetization-prepared rapid gradient echo imaging (MPRAGE)에서 inversion-recovery spoiled gradient recalled (IR-SPGR)로의 시퀀스의 변화는 평균적으로 -1.63%의 변화를 보여, 제조사 간의 차이와 유사했다. 반면 각 제조사 내에서 기기의 업그레이드는 0.5% 이하의 변화를 보였고, 소프트웨어 업그레이드는 유의한 변화를 초래하지 않았다.

같은 3T 내에서의 MR 기기변경이나 업그레이드의 잠재적 효과에 대해서는, 유사한 결과가 예측되나 좀 더 정확한 효과를 알기 위해, 향후 연구가 필요하다. 이런 관점에서 대부분의 임상용 소프트웨어에서는 정상군의 데이터를 제공하고, 정상 데이터로부터의 벗어난 정도를 제공함으로써 용적측정에서의 오차 문제를 해결하려고 하고 있다. 현재 대부분의 임상용 소프트웨어는 각기 서로 다른 정상인들의 데이터를 이용한 정상 범위를 제공하고 있다.

하지만, 성별, 국가 및 인종에 따라 뇌의 전체 용적 및 부분적인 용적이 다르다는 해부학적 보고가 있고, 최근 MS metrix를 이용한 연구에서는 2%~3% 정도의 정상 브라질 대조군이 유럽 기준으로 5 percentiles 이하로 판정되어, 지역특성에 맞는 정상군의 필요성을 보여주었다(28). 해외에서는 2000명 이상의 환자 및 정상군 데이터가 공공 데이터로 구축되어 있으나, 현재 국내에서는 연구자들이나 혹은 소프트웨어 개발자가 공개적으로 쓸 수 있는 정상군 데이터가 존재하지 않아서 개발 단계에서 재현성 검증 연구가 쉽지 않고, 개발 후에도 재현성 면에서 실패 원인이 될 수 있다.

한편, 병태생리적인 제한점은 근본적으로 뇌용적 분석이 진단하고자 하는 질환의 절대적 기준이 아니라는 점이다. 뇌용적의 감소는 어느 특정 질환에 특이적이지 않은데, 퇴행성 질환 외에도 외상, 알코올 섭취 등 다양한 원인에 의해 뇌용적이 감소할 수 있다. 그러므로 뇌 위축이 중심 병리 중 하나인 뇌 질환을 평가할 때도 다양한 다른 형태학적 표지자들을(sulcal span, sulcal shape, mean curvature index) 모두 이용하여 연구를 해야 하며, 이를 고려할 때 특정 부위의 위축이라는 제한된 지표로 질병의 진단이나 심각도를 판단하는 것은 위양성과 위음성의 위험성을 내포한다.

또 병태생리적으로 퇴행성 뇌 질환의 뇌 위축은 시간의 축에 따라 뇌용적이 변화하는 정도가 특정 시기의 용적 값보다 더욱 중요한 의미를 지닌다. 따라서 추적관찰 시의 변화가 정상적으로 어느 정도인지, 뇌용적 변화 평가를 위한 추적검사의 적절한 시간 간격은 어느 정도인지, 추적검

사와 MR 기기 변화/업그레이드와는 어떻게 조화시킬지 등의 다양한 문제의 해결이 필요하다.

## 대처방안에 대한 전문가 권고안

### 최적의 정상 표준치 확립을 위한 공동의 노력

기존의 뇌용적 분석 임상용 소프트웨어는 서로 다른 정상군에서 얻어진 표준을 사용하고 있다. 이는 나이, 성별, 지역, 계층, 인종을 고려하지 않은 데이터로 국내 임상 적용을 위해서는 국내 정상인 빅데이터를 확보하고 이를 실제로 소프트웨어에 표준데이터로 제공할 수 있는 공동의 노력이 필요하다. 또한 기관마다 서로 다른 환자군을 진료하는 것을 감안하여, 임상혈액검사와 유사하게, 기관별로 최적화된 정상치를 설정하는 것도 필요하다.

### 프로토콜 및 시퀀스의 표준화

뇌용적 분석 소프트웨어의 신뢰도를 위해 MRI 영상 검사 프로토콜을 표준화하는 것이 가장 중요한 부분이나, 흔히 같은 기기에서 촬영하는 경우 간과하기 쉽다. 또한, 3T와 1.5T 간의 차이, 사용한 head coil의 array 수 등 영상 품질 간의 차이를 최소화할 수 있는 방법을 연구를 통해 개발하여야 한다. 임상현장에서 환자 추적검사, 동일 기관 혹은 다기관 연구 등과 같은 상황에서 최적화된 표준영상시퀀스를 이용한 프로토콜을 이용할 수 있도록 주의를 기울여야 한다.

### 분석 전, 분석 후 전문가 모니터링의 필요성

연구용 프로그램에서의 실패율은 보고에 따르면 6%~8% 정도이고(34), 국내 병원에서 조사한 결과도 최대 7%에 이른다. 분석 실패의 원인은 잘못된 영상데이터 외에도, 영상 획득 시 심한 움직임 등으로 인한 환자에서 기인한 문제로 용적이 잘못 계산될 수 있다. 따라서 분석 전후로 전문가가 주의하여 확인하여야 할 부분은 MR 시스템 자체의 문제, 환자의 문제, 영상 인공물의 유무 평가 등이다. 또한 반드시 소프트웨어 분석 시 오류 유무를 평가하는 것도 필요하다. 앞으로는 이 부분도 인공지능을 이용한 시도를 통해 개선될 수 있을 것으로 생각한다.

### 영상 획득, 분석, 판독 최적화를 위한 영상의학과 전문의의 노력

임상용 소프트웨어의 설치와 분석 및 결과 판독 및 임상 적용의 전 영역에 걸쳐 영상의학과 전문의의 감독이 필요하며, 이로써 최종 판독과 진단의 정확도를 향상시켜야 한다.

현재 소프트웨어 기술의 급속한 발달과 인공지능의 적용 등으로 인하여 소프트웨어 개발자와 최종적으로 이용하는 의료진 간의 지식의 간극이 커지면서 소프트웨어가 오용/남용될 가능성이 점점 더 우려된다. 따라서, 영상의학과 전문의는 조정자로서 전문가 집단 간의 긴밀한 상호관계와 지식 교환을 통해 이러한 간극을 좁히는 노력을 할 필요가 있다. 또 소프트웨어 개발자는 의료진과의 공동 연구를 통해 개발 소프트웨어의 적절한 오차범위 및 재현성 자료를 공유하고, 진료 현장의 의료진에게 이를 전달해야 할 의무가 있다.

## 결론

결론적으로, 다양한 종류의 임상용 뇌용적 분석 소프트웨어의 이용은 좀 더 광범위해질 것이며, 이는 인공지능과 빅데이터 적용으로 더욱 급속하게 진행될 것이다. 따라서 소프트웨어의 올바른 임상 적용을 위해 각 소프트웨어들의 장단점을 미리 숙지하고 소프트웨어의 사용으로 발생할 수 있는 문제점을 영상의학과 전문의가 충분히 미리 인지해야 한다.

## Supplementary Materials

The online-only Data Supplement is available with this article at <http://dx.doi.org/10.3348/jksr.2020.0174>.

## Author Contributions

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

## Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

## Funding

This study was supported by a grant by 2017 Clinical Practice Guideline Research Fund by Korean Society of Radiology.

## Acknowledgments

KSNR Taskforce on degenerative disease members are as following in alphabetical order: Sung Tae Kim, MD (Sungkyunkwan University School of Medicine), Seul Kee Kim (Chonnam National University Medical School), Ji Hoon Kim, MD (Seoul National University College of Medicine), Mina Park (Yonsei University College of Medicine), Sun Won Park, MD (Seoul National University College of Medicine), Sung Jun Ahn (Yonsei University College of Medicine), Kyung Mi Lee (Kyung Hee University College of Medicine), Geon-Ho Jahng (Kyung Hee University Hospital at Gangdong, College of Medicine), Jinhee Jang (College of Medicine, Catholic University of Korea), Chi-Hoon Choi (Chungbuk National University College of Medicine), Hyun Seok Choi (Yonsei University College of Medicine).

The authors sincerely thank Sam-Soo Kim Kim, MD and Kook-Jin Ahn MD for their most valuable advice and support in developing recommendations.

## REFERENCES

1. Jack CR Jr. Alzheimer disease: new concepts on its neurobiology and the clinical role imaging will play. *Radiology* 2012;263:344-361
2. Jack CR Jr, Albert MS, Knopman DS, McKhann GM, Sperling RA, Carrillo MC, et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:257-262
3. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:263-269
4. 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
5. 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
6. 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
  7. Tanpitukpongse TP, Mazurowski MA, Ikkena 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
  8. 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
  9. Louis S, Morita-Sherman M, Jones S, Vegh D, Bingaman W, Blumcke I, et al. Hippocampal sclerosis detection with NeuroQuant compared with neuroradiologists. *AJNR Am J Neuroradiol* 2020;41:591-597
  10. Guo C, Ferreira D, Fink K, Westman E, Granberg T. Repeatability and reproducibility of FreeSurfer, FSL-SIENAX and SPM brain volumetric measurements and the effect of lesion filling in multiple sclerosis. *Eur Radiol* 2019;29:1355-1364
  11. Storelli L, Rocca MA, Pagani E, Van Hecke W, Horsfield MA, De Stefano N, et al. Measurement of whole-brain and gray matter atrophy in multiple sclerosis: assessment with MR imaging. *Radiology* 2018;288:554-564
  12. Ross DE, Ochs AL, Tate DF, Tokac U, Seabaugh J, Abildskov TJ, et al. High correlations between MRI brain volume measurements based on NeuroQuant<sup>®</sup> and FreeSurfer. *Psychiatry Res Neuroimaging* 2018;278:69-76
  13. Chung J, Kim H, Moon Y, Moon WJ. Comparison of vendor-provided volumetry software and NeuroQuant using 3D T1-weighted images in subjects with cognitive impairment: how large is the inter-method discrepancy? *Investig Magn Reson Imaging* 2020;24:76-84
  14. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *Neuroimage* 2012;62:782-790
  15. Fischl B. FreeSurfer. *Neuroimage* 2012;62:774-781
  16. Ashburner J, Friston KJ. Unified segmentation. *Neuroimage* 2005;26:839-851
  17. Petrella JR. Neuroimaging and the search for a cure for Alzheimer disease. *Radiology* 2013;269:671-691
  18. Ochs AL, Ross DE, Zannoni MD, Abildskov TJ, Bigler ED; Alzheimer's Disease Neuroimaging Initiative. Comparison of automated brain volume measures obtained with NeuroQuant<sup>®</sup> and FreeSurfer. *J Neuroimaging* 2015;25:721-727
  19. Reid MW, Hannemann NP, York GE, Ritter JL, Kini JA, Lewis JD, et al. Comparing two processing pipelines to measure subcortical and cortical volumes in patients with and without mild traumatic brain injury. *J Neuroimaging* 2017;27:365-371
  20. Persson K, Barca ML, Cavallin L, Brækhus A, Knapskog AB, Selbæk G, et al. Comparison of automated volumetry of the hippocampus using NeuroQuant<sup>®</sup> and visual assessment of the medial temporal lobe in Alzheimer's disease. *Acta Radiol* 2018;59:997-1001
  21. 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
  22. Persson K, Selbæk G, Brækhus A, Beyer M, Barca M, Engedal K. Fully automated structural MRI of the brain in clinical dementia workup. *Acta Radiol* 2017;58:740-747
  23. Ross DE, Seabaugh J, Cooper L, Seabaugh J. NeuroQuant<sup>®</sup> and NeuroGage<sup>®</sup> reveal effects of traumatic brain injury on brain volume. *Brain Inj* 2018;32:1437-1441
  24. Eggins PS, Hatton SN, Hermens DF, Hickie IB, Lagopoulos J. Subcortical volumetric differences between clinical stages of young people with affective and psychotic disorders. *Psychiatry Res Neuroimaging* 2018; 271:8-16
  25. Ahdidan J, Raji CA, DeVoe EA, Mathis J, Noe KØ, Rimestad J, et al. Quantitative neuroimaging software for clinical assessment of hippocampal volumes on MR imaging. *J Alzheimers Dis* 2016;49:723-732
  26. Fragoso YD, Wille PR, Abreu M, Brooks JBB, Dias RM, Duarte JA, et al. Correlation of clinical findings and brain volume data in multiple sclerosis. *J Clin Neurosci* 2017;44:155-157
  27. Steenwijk MD, Amiri H, Schoonheim MM, de Sitter A, Barkhof F, Pouwels PJW, et al. Agreement of MSmetrix with established methods for measuring cross-sectional and longitudinal brain atrophy. *Neuroimage Clin* 2017;15:843-853
  28. Finkelsztejn A, Fragoso YD, Bastos EA, Duarte JA, Varela JS, Houbrechts R, et al. Intercontinental validation of brain volume measurements using MSmetrix. *Neuroradiol J* 2018;31:147-149
  29. Niemantsverdriet E, Ribbens A, Bastin C, Benoit F, Bergmans B, Bier JC, et al. A retrospective Belgian multi-center MRI biomarker study in Alzheimer's disease (REMEMBER). *J Alzheimers Dis* 2018;63:1509-1522

30. Lee JS, Kim C, Shin JH, Cho H, Shin DS, Kim N, et al. Machine learning-based individual assessment of cortical atrophy pattern in Alzheimer's disease spectrum: development of the classifier and longitudinal evaluation. *Sci Rep* 2018;8:4161
31. Biberacher V, Schmidt P, Keshavan A, Boucard CC, Righart R, Sämann P, et al. Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. *Neuroimage* 2016;142:188-197
32. Park JE, Park SY, Kim HJ, Kim HS. Reproducibility and generalizability in radiomics modeling: possible strategies in radiologic and statistical perspectives. *Korean J Radiol* 2019;20:1124-1137
33. Fellhauer I, Zöllner FG, Schröder J, Degen C, Kong L, Essig M, et al. Comparison of automated brain segmentation using a brain phantom and patients with early Alzheimer's dementia or mild cognitive impairment. *Psychiatry Res* 2015;233:299-305
34. Mulder ER, de Jong RA, Knol DL, van Schijndel RA, Cover KS, Visser PJ, et al. Hippocampal volume change measurement: quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST. *Neuroimage* 2014;92:169-181
35. Cover KS, van Schijndel RA, Versteeg A, Leung KK, Mulder ER, Jong RA, et al. Reproducibility of hippocampal atrophy rates measured with manual, FreeSurfer, AdaBoost, FSL/FIRST and the MAPS-HBSI methods in Alzheimer's disease. *Psychiatry Res Neuroimaging* 2016;252:26-35
36. Nestor SM, Gibson E, Gao FQ, Kiss A, Black SE; Alzheimer's Disease Neuroimaging Initiative. A direct morphometric comparison of five labeling protocols for multi-atlas driven automatic segmentation of the hippocampus in Alzheimer's disease. *Neuroimage* 2013;66:50-70
37. Yim Y, Lee JY, Oh SW, Chung MS, Park JE, Moon Y, et al. Comparison of automated brain volume measures by NeuroQuant vs. Freesurfer in patients with mild cognitive impairment: effect of slice thickness. *Yonsei Med J* 2021;62:255-261
38. Lee JB, Lee JY, Oh SW, Chung MS, Park JE, Moon Y, et al. Evaluation of reproducibility of brain volumetry between Inbrain and FreeSurfer. *J Clin Neurol* 2021;17:307-316
39. 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 Radiol* 2021;22:405-414
40. 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
41. Azab M, Carone M, Ying SH, Yousem DM. Mesial temporal sclerosis: accuracy of NeuroQuant versus neuroradiologist. *AJNR Am J Neuroradiol* 2015;36:1400-1406
42. Potvin O, Duchesne S. IC-P-120: normative structural neuroimaging data for age and sex using freesurfer. *Alzheimer's & Dementia* 2015;11:82
43. Lee H, Nakamura K, Narayanan S, Brown RA, Arnold DL; Alzheimer's Disease Neuroimaging Initiative. Estimating and accounting for the effect of MRI scanner changes on longitudinal whole-brain volume change measurements. *Neuroimage* 2019;184:555-565

## 뇌 자기공명영상 뇌용적 분석 소프트웨어의 임상적 적용에 대한 전문가 의견과 권고안

이지영<sup>1</sup> · 박지은<sup>2</sup> · 정미선<sup>3</sup> · 오세원<sup>4</sup> · 문원진<sup>5\*</sup>;  
대한신경두경부영상의학회 및 산하 퇴행성신경질환연구회

치매를 비롯한 퇴행성 신경 질환의 초기 진단에 자기공명영상을 이용한 뇌 위축 평가와 정량적 용적 분석이 중요하다. 뇌 위축의 시각적 평가는 주관적으로 평가자에 따라 다른 결과를 보여주기 때문에, 객관적인 결과를 제공하면서 임상 적용도 가능한 소프트웨어의 수요와 개발이 늘어나고 있다. 이러한 임상용 소프트웨어의 실제 임상 적용은 영상 검사의 표준화가 선행되어야 하고, 개발된 소프트웨어의 검증이 반드시 필요하다. 따라서 대한신경두경부영상의학회는 뇌용적 분석 임상용 소프트웨어의 임상적 활용에 대한 의견을 제시하기 위해 전문위원회를 구성하고 현재까지 발표된 연구를 정리하였다. 그리고, 정량화 분석을 위한 영상 검사의 표준화 및 소프트웨어의 임상 적용에 대한 전문가 의견을 제시하기 위하여 공동 작업을 수행하였다. 본 종설에서는 뇌 자기공명영상의 정량화 분석의 필요성 및 배경, 정량화 분석을 위한 임상용 소프트웨어의 소개 및 기존의 표준품(reference standard)과의 진단능 비교, 영상 획득의 표준화, 분석 및 평가의 표준화, 소프트웨어의 임상 적용에 대한 전문가 의견, 제한점 및 대처 방법 등 대한신경두경부영상의학회의 전문가 권고안을 소개하는 것이 목적이다.

<sup>1</sup>한양대학교 의과대학 한양대학교병원 영상의학과,

<sup>2</sup>울산대학교 의과대학 서울아산병원 영상의학과,

<sup>3</sup>중앙대학교 의과대학 중앙대병원 영상의학과,

<sup>4</sup>가톨릭대학교 의과대학 은평성모병원 영상의학과,

<sup>5</sup>건국대학교 의학전문대학원 건국대학교병원 영상의학과