Deep learning-based diagnosis of Alzheimer’s disease using brain magnetic resonance images: an empirical study

Jun Sung Kim1,2,14, Ji Won Han2,3,14, Jong Bin Bae2, Dong Gyu Moon2, Jin Shin2, Juhee Eliana Kong2, Hyungji Lee2, Hee Won Yang3, Eunji Lim2, Jun Yup Kim4, Leonard Sunwoo2,7,8, Se Jin Cho2,7,8, Dongsoo Lee9, Injoong Kim10, Sang Won Ha11, Min Ju Kang11, Chong Hyun Suh12, Woo Hyun Shim12, Sang Joon Kim12 & Ki Woong Kim1,2,3,13*

The limited accessibility of medical specialists for Alzheimer’s disease (AD) can make obtaining an accurate diagnosis in a timely manner challenging and may influence prognosis. We investigated whether VUNO Med-DeepBrain AD (DBAD) using a deep learning algorithm can be employed as a decision support service for the diagnosis of AD. This study included 98 elderly participants aged 60 years or older who visited the Seoul Asan Medical Center and the Korea Veterans Health Service. We administered a standard diagnostic assessment for diagnosing AD. DBAD and three panels of medical experts (ME) diagnosed participants with normal cognition (NC) or AD using T1-weighted magnetic resonance imaging. The accuracy (87.1% for DBAD and 84.3% for ME), sensitivity (93.3% for DBAD and 80.0% for ME), and specificity (85.5% for DBAD and 85.5% for ME) for diagnosing AD were comparable; however, DBAD showed a higher trend in every analysis than ME diagnosis. DBAD may support the clinical decisions of physicians who are not specialized in AD; this may enhance the accessibility of AD diagnosis and treatment.

The number of people with Alzheimer’s disease (AD) is estimated to be more than 50 million worldwide1 and is expected to increase more than three-fold within the next 30 years2. However, in many regions of the world, it is a challenge to obtain an accurate diagnosis in a timely manner owing to limited accessibility to AD medical specialists. For example, hospitals that offer specialized diagnosis and management of AD are usually concentrated in urban areas, although AD is more prevalent in rural areas3. This rural–urban inequality in accessibility to specialists may contribute to disparities in health outcomes of people with AD between rural and urban areas4. Structural neuroimaging, such as brain magnetic resonance imaging (MRI), is an essential test for diagnosing AD and monitoring its course5. According to the Organization for Economic Co-operation and Development, the number of MRI units is constantly increasing worldwide6. However, in many regions, the results of brain MRI cannot be interpreted in a timely and accurate manner owing to a lack of AD specialists. Moreover, even AD specialists...
specialists cannot identify the early changes in AD that occur on brain MRI with the naked eye\cite{2,9}. A technology that can accurately identify people with AD from the prodromal or early dementia stages using brain MRI would increase the diagnostic rate of AD and advance the time of diagnosis of AD, thus reducing the health outcome disparities between regions by supporting the clinical decision of physicians who are not specialized in AD.

Deep learning (DL) using convolutional neural networks (CNN) has been proposed as a promising tool for supporting clinical decisions on digital brain images of AD\cite{10-15}. In many previous studies, three-dimensional (3D) brain MR images were employed as input data in DL algorithms for diagnosing AD\cite{10-14,16-18}. However, 3D brain MRI is not available in most clinical settings because 3D MRI sequences have longer acquisition times than two-dimensional (2D) MRI sequences\cite{19}, which significantly increase computational burden, storage, and cost\cite{20,21}. Thus, the previously developed DL algorithms for diagnosing AD using 3D MRI may be difficult to apply to most brain MRI scans obtained in typical clinical settings. Therefore, we developed VUNO Med-DeepBrain AD, version 1.0.0 (DBAD; VUNO Inc., Seoul, Korea), which is the first convolutional neural network-based model for diagnosing AD using 2D brain MR images as input data, and demonstrated its diagnostic accuracy for AD to be excellent in both Caucasians and Asians\cite{22}.

This study aimed to investigate whether VUNO Med-DeepBrain AD using 2D brain MRI can be employed as a decision support service for AD diagnoses in hospitals, by comparing the diagnostic performance of VUNO Med-DeepBrain AD with that of AD medical specialists working at referral hospitals. If VUNO Med-DeepBrain AD would be found to be as accurate as the decision of medical specialists, it can be easily implemented in real clinical settings because it has fast processing speed and uses 2D images that are commonly employed in usual clinical settings. Introduction of VUNO Med-DeepBrain AD may contribute to the early diagnosis of AD not only in memory clinics but also in any medical settings that uses brain MRI.

**Methods**

**Participants.** We enrolled 100 older adults aged 60 or older (34 with a history of AD and 66 without a history of AD) from the visitors to the Seoul Asian Medical Center (AMC) and the Korea Veterans Health Service (KVHS). Then we administered the standardized diagnostic assessment for AD to 98 participants of them at the Seoul National University Bundang Hospital (SNUBH) after excluding two participants who refused the standard diagnostic assessment. All participants and/or their legal guardians provided a written informed consent to participate in this study. We obtained ethics approval from the Institutional Review Board of AMC, KVHS and SNUBH. All experiments were performed in accordance with relevant guidelines and regulations.

**Acquisition and preprocessing of brain MRI.** We acquired 3D T1-weighted MR images in the Digital Imaging and Communications in Medicine format using a 3.0 Tesla InVivo scanner (Philips Medical Systems; Eindhoven, NL) at the AMC and a 3.0 Tesla Magnetom Skyra scanner (Siemens Healthineers, Erlangen, Germany) or a 3.0 Tesla Magnetom Vida scanner (Siemens Healthineers) at the KVHS. The parameters were as follows: repetition time = 9.6 ms, echo time = 4.6 ms, flip angle = 8°, field of view = 224 × 224 mm², slice thickness = 1 mm with a no gap, and matrix size = 224 × 224 mm² in the InViva scanner; repetition time = 1900 ms, echo time = 2.9 ms, flip angle = 9°, field of view = 230 × 230 mm², slice thickness = 1 mm, and matrix size = 256 × 256 mm² in the Magnetom Skyra scanner; and repetition time = 1900 ms, echo time = 2.9 ms, flip angle = 9°, field of view = 230 × 230 mm², slice thickness = 1 mm, and matrix size = 256 × 256 mm² in the Magnetom Vida scanner.

We resampled the input images into a grid of 256 x 256 x 256 voxels with an isotropic spatial resolution of 1 × 1 × 1 mm³ using the mri_convert routine in FreeSurfer\cite{23}. Then we extracted 2D coronal slices around the medial temporal lobe through two stages of rigid transformation using DBAD. In the first stage of rigid transformation, DBAD matched the position of the input images to a Korean normal elderly template (KNE) constructed from a cognitively normal elderly population\cite{24}. Using the template-registered input image, DBAD extracted the brain parenchyma using the custom brain extraction algorithm, which is based on a 3D U-Net generated by the Brain Extraction Tool in the FMRIB Software Library\cite{25}. In the second stage of rigid transformation, DBAD registered the output images of the first stage to a skull-stripped version of the KNE and extracted 256 2D coronal slices. Based on the selection criteria of the medial temporal lobe atrophy visual rating scale\cite{26}, DBAD selected 30 consecutive 2D coronal slices, starting from the corpus of the hippocampus, that were used as the input images in the DBAD CNN-based model for diagnosing AD. DBAD applied min-max normalization to bound the values of the images between zero and one in all slices\cite{27}.

**MRI-based diagnosis.** The DBAD fed 30 2D coronal slices of medial temporal lobe with age and sex of each participant into the DBAD CNN-based model for diagnosing AD. The DBAD CNN-based model for diagnosing AD uses F Inception-V4 as its backbone and extracts various features that include structural and textural information of the input images. The DBAD CNN-based model for diagnosing AD concatenated age, sex, and the location information of the coronal slices (slice number) and entered them into a fully connected network that calculates the probability of AD of each slice. The DBAD CNN-based model for diagnosing AD averaged the probabilities of AD of the slices to calculate a final score that represents the subject's probability of having AD (DBAD score, score ranges from 0 to 1). Then we classified the participants with the DBAD score of ≥ 0.38 as AD (DBAD-AD), and those with the DBAD score of < 0.38 as normal (DBAD-CT) based on our previous study\cite{28} (Fig. 1). Deep learning model was implemented using Pytorch (v0.4.1) and it was conducted using NVIDIA Geforce GTX 1080 Ti GPU.

Three medical experts (two neuroradiologists and one neurologist) also classified each participant into AD (ME-AD) and normal (ME-CT) groups using full volumetric, T1-weighted MRI scans as well as participant’s age and sex. Each medical expert was blinded to the decisions of other medical experts (ME) and DBAD. We considered the diagnosis that two or more ME agreed upon the consensus diagnosis of the panel.
Acquisition of $^{18}$F-florbetaben amyloid brain PET scans and determination of amyloid positivity. Brain $^{18}$F-florbetaben amyloid PET images were obtained using a Discovery VCT scanner (General Electric Medical Systems; Milwaukee, WI, USA). We injected 8.1 mCi (300 MBq) of $^{18}$F-florbetaben (Neuraceq, Piramal, Mumbai, India) as a slow single intravenous bolus (6 s/mL) in a total volume of up to 10 mL and obtained 20-min PET images comprising four 5-min dynamic frames after a 90-min uptake period. Trained radiologists with expertise in nuclear medicine determined amyloid beta peptide (Aβ)-positivity based on the brain amyloid plaque load (BAPL) score. The BAPL score is a predefined three-grade scoring system wherein measurements are made by the physician according to the visual assessment of the participant's amyloid deposition in the brain using Neuraceq. BAPL scores of 1 (BAPL 1), 2 (BAPL 2), and 3 (BAPL 3) indicate no Aβ load, minor Aβ load, and significant Aβ load, respectively. Therefore, BAPL 1 indicates an Aβ-negative status, whereas BAPL 2 and BAPL 3 indicate an Aβ-positive status.

Standard diagnosis. Geriatric neuropsychiatrists with expertise in dementia research (not participating in the MRI-based diagnoses) administered a face-to-face standardized diagnostic interview as well as physical and neurological examinations using the Korean version of the Consortium to Establish a Registry for Alzheimer’s Disease Assessment Packet Clinical Assessment Battery (CERAD-K-C) to diagnose cognitive disorders. Laboratory tests, including complete blood counts, chemistry profiles, and serological tests for syphilis, were performed for each participant. Research neuropsychologists, geriatric psychiatrists, and trained research nurses blinded to the MRI-based diagnoses administered the CERAD-K-N, which consists of the following neuropsychological tests: Verbal Fluency Test, 15-item Boston Naming Test, Mini-Mental State Examination (MMSE), Word List Memory Test, Constructional Praxis Test, Word List Recall Test, Word List Recognition Test, Constructional Recall Test, Trail Making Test A/B, Digit Span Test, and Frontal Assessment Battery.

We then determined the standard diagnosis and Clinical Dementia Rating (CDR) of each participant at the diagnostic conference in which four research geriatric psychiatrists participated. We diagnosed mild cognitive impairment (MCI) according to the consensus criteria from the International Working Group on MCI; dementia according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) diagnostic criteria; and AD according to the National Institute on Aging-Alzheimer’s Association workgroups (NIA-AA) criteria. All AD patients were Aβ-positive on $^{18}$F-florbetaben brain PET scan and met the NIA-AA criteria of preclinical AD and MCI or dementia due to AD.

Statistical analysis. We compared the continuous and categorical variables between groups using Mann–Whitney U-test and Pearson chi-square test, respectively. We then compared the diagnostic classification between DBAD and panel of medical experts (ME). We estimated sensitivity, specificity and accuracy of the diagnoses made by the DBAD and the ME using the receiver operating characteristic (ROC) curve analysis. Then, we compared the classification performance of the DBAD and the ME using McNemar’s test.

All analyses were performed using SPSS for Windows (version 25.0; IBM Co., Armonk, NY, USA) and MedCalc version 16.4.3 (MedCalc Software, Mariakerke, Belgium). We considered two-sided p-values < 0.05 to indicate statistical significance.

Results
The demographic characteristics of the participants are summarized in Table 1. The ME-AD and DBAD-AD groups were older ($p<0.001$) and less educated than the ME-CT ($p=0.016$) and DBAD-CT ($p<0.001$) groups, respectively (Table 1).
As summarized in Table 2, the ME classified 36 participants as ME-AD and the rest as ME-CT. The decisions of the three medical experts completely matched for 77 participants (26 ME-AD and 51 ME-CT), and their κ agreement was 0.694 ($p < 0.001$). Among the 36 ME-AD participants, 15 were diagnosed with AD in the standard diagnostic assessment (12 dementia due to AD, one MCI due to AD, and two preclinical AD). Among the 21 ME-AD participants who were not diagnosed with AD in the standard diagnostic assessment, 6 were diagnosed with other types of dementia (4 subcortical vascular dementia and 2 with suspected non-AD pathophysiology [SNAP]), 7 with Aβ-negative MCI, and 8 were Aβ-negative and cognitively normal. Among the 62 ME-CT participants, 47 were Aβ-negative and cognitively normal according to the standard diagnostic assessment. In contrast, among the remaining 15 participants, 9 were diagnosed with AD (three with dementia due to AD, three with MCI due to AD, and three with preclinical AD), 1 with SNAP, and 5 with Aβ-negative MCI.

The DBAD classified 37 participants as DBAD-AD and the rest as DBAD-CT. Among the 37 DBAD-AD participants, 17 were finally diagnosed with AD in the standard diagnostic assessment (14 with dementia due to AD, 2 with MCI due to AD, and 1 with preclinical AD). Among the 20 DBAD-AD participants who were not diagnosed with AD in the standard diagnostic assessment, 6 were diagnosed with other types of dementia (4 subcortical vascular dementia, 2 SNAP), 6 with Aβ-negative MCI, and 8 were Aβ-negative and cognitively normal. Among the 61 DBAD-CT participants, 47 were Aβ-negative or cognitively normal according to the standard diagnostic assessment. In contrast, among the remaining 14 DBAD-CT participants, 7 were diagnosed with AD (1 dementia due to AD, 2 MCI due to AD, and 4 preclinical AD), 1 with SNAP, and 6 with Aβ-negative MCI (Table 2).

As summarized in Table 3, the accuracy for classifying Aβ-positive AD patients from Aβ-negative cognitively normal controls was comparable between DBAD and the ME consensus diagnosis. The specificity of classifying Aβ-negative cognitively normal controls as normal was the same for DBAD and ME. The sensitivity of DBAD in diagnosing Aβ-positive AD patients was higher than that of ME at all levels of the definition of AD, although the differences were not significant. However, when each ME diagnosis was analyzed separately, the accuracy of DBAD was higher than that of one ME ($p = 0.031$ for dementia due to AD, $p = 0.009$ for dementia/MCI due to AD, and $p = 0.016$ for dementia/MCI/preclinical due to AD).
As summarized in Table 4, the accuracy of distinguishing Aβ-positive AD patients from Aβ-negative non-AD patients and the specificity for classifying Aβ-negative participants as non-AD were also comparable between the DBAD and ME groups. The sensitivity of DBAD in diagnosing Aβ-positive AD patients was higher than that of ME at all levels of the definition of AD, but the differences were not significant (Table 4). The DBAD accuracy of distinguishing Aβ-positive AD patients from Aβ-negative cognitively normal controls was lower than that of distinguishing Aβ-positive AD patients from Aβ-negative cognitively normal controls, because the specificity for classifying Aβ-negative participants as non-AD was lower than that for classifying Aβ-negative cognitively normal controls as normal.

**Discussion**

This study demonstrated that DBAD can support the decision on the diagnosis of AD in the hospitals where medical specialists on AD are not available by showing that the diagnostic accuracy for AD of DBAD was comparable to that of the medical experts on AD who are working at the referral hospitals.
Medial temporal lobe atrophy (MTA) is now considered a valid diagnostic marker of AD at the MCI stage\(^\text{38}\), and the MTA rating is widely used in clinical practice to determine the presence of AD-related neurodegeneration\(^\text{22}\). The DBAD uses 30 coronal slices starting from the corpus of the hippocampus based on the Scheltens score for MTA ratings\(^\text{22,26}\), whereas an ME panel can use additional information such as cortical atrophy patterns, ventricular enlargement, or small vessel disease through the full volume of T1-weighted MRI. Despite this disparity in the information provided, DBAD demonstrated an expert level of diagnostic accuracy and was superior to individual ME diagnosis.

The diagnostic accuracy and sensitivity of both DBAD and the ME were better for dementia due to AD than for MCI due to AD and/or preclinical AD. DBAD is an algorithm developed and validated using a divided dataset consisting of normal controls and dementia due to mild AD (CDR 0.5 or 1) for typical normal and AD classification\(^\text{22}\). Therefore, the diagnostic performance of DBAD may be lower when diagnosing preclinical AD or MCI due to AD than when diagnosing dementia due to AD. MTA is less pronounced in MCI and preclinical AD than in dementia due to AD\(^\text{37}\). The University of California-Los Angeles Alzheimer’s Disease Research Center has studied the hippocampal volume loss between MCI due to AD and dementia due to AD using 20 follow-up MCI participants and revealed that the annual atrophy rate for those who remained in the MCI stage due to AD was 2.8%, while that for those who developed dementia due to AD was 3.7%\(^\text{38}\). Moreover, the Mayo Clinic AD Research Center/AD Patient Registry also studied the hippocampal volume loss of normal controls, patients with MCI, and those with probable AD in a sample of 129 participants and found that the mean annualized rates of hippocampal volume loss were 1.73% for normal controls, 2.5% for MCI, and 3.5% for probable AD\(^\text{39}\). This body of literature implies that MTA is a later event in AD progression\(^\text{40}\), which may also explain the lower diagnostic performance of DBAD when including preclinical and MCI stages of AD than when only including dementia due to AD. DBAD exploits image slices that represent MTA in the coronal view from 3D T1-weighted MRI to diagnose AD. However, DL using MRI not only helps interpret and diagnose the volume reduction or atrophy of the brain but also provides a comprehensive reflection of other sensitive features (e.g., MRI texture\(^\text{41}\)) related to AD that DBAD may have analyzed during diagnosis.

Nevertheless, even after including preclinical AD and MCI due to AD, the diagnostic performance of DBAD was comparable to that of ME. This supports the promising role of DBAD as a diagnostic assistant tool for AD, even in the early stages of AD in a clinical setting in the absence of an ME. In addition, the diagnostic accuracy, sensitivity, and specificity of DBAD in diagnosing AD showed a similar level of performance to that in the previous paper for algorithm development\(^\text{22}\). This study may also serve as an independent validation study for the diagnostic accuracy, sensitivity, and specificity of DBAD.

Various studies on DL algorithms to diagnose AD have been conducted\(^\text{12,13,16–18}\); however, few algorithms have been validated by comparisons of DL algorithm performance with that of medical doctors. Qiu et al.\(^\text{42}\) created an algorithm for diagnosing AD using DL with the 417 Alzheimer’s Disease Neuroimaging Initiative dataset. They compared the diagnostic performance of the developed algorithm and that of neurologists after they were provided with full volumetric T1-weighted MRI scans, age, sex, and MMSE scores of the participants. In line with our study, their DL model (accuracy: 0.834 ± 0.020) outperformed the neurologists (accuracy: 0.823 ± 0.094). In addition, a study by Nagendran\(^\text{43}\) gathered the research that compared the diagnostic performance of DL algorithms and clinicians, based on medical imaging for diseases such as cancer, cataract, and colon polyps. Among a total of 77 papers comparing DL and clinicians’ diagnoses, 30% studies showed that DL performance was superior to clinicians,’ 17% said DL was comparable or better, 32% said DL was comparable, 18% said DL was able to help a clinician perform better, and 3% said DL was not superior. These previous studies suggest the usefulness of the DL algorithm as a supporting tool in clinical settings.

The current study has some limitations. First, the number of ME included was limited, and a larger number of ME are needed for further validation studies. Second, although DBAD showed comparable performance to that of ME, we would consider incorporating other AD signature regions, such as the precuneus\(^\text{44,45}\) and posterior cingulate cortex\(^\text{46,47}\) to the DBAD algorithm to enhance the diagnostic accuracy even in the preclinical AD stage. However, in this case, we may have to sacrifice the current strength of DBAD processing speed.

In this study, we validated the diagnostic performance of AD classification CNN-based algorithm by comparing it with MRI diagnosis by medical experts. The VUNO Med-DeepBrain AD may support the clinical decision of the physicians who are not specialized in AD, which may enhance the accessibility of AD diagnosis and treatment.

Data availability
Data are not publicly available because they contain sensitive participant information. Individual, de-identified, participant data that underlie the results reported in this article may be made available to qualified researchers upon reasonable request. Proposals should be directed to K.W.K. to gain access.

Received: 16 June 2022; Accepted: 20 October 2022
Published online: 26 October 2022

References
1. Patterson, C. World Alzheimer Report 2018. (2018).
2. Hebert, L. E., Weuve, J., Scherr, P. A. & Evans, D. A. Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. Neurology 80, 1778–1783. https://doi.org/10.1212/WNL.0b013e31828726f5 (2013).
3. Bae, J. B. et al. Incidence of and risk factors for Alzheimer’s disease and mild cognitive impairment in Korean elderly. Dement. Geriatr. Cogn. Disord. 39, 105–115. https://doi.org/10.1159/000366555 (2015).
4. Johnston, K. J., Wen, H. & Joynt Maddox, K. E. Lack of access to specialists associated with mortality and preventable hospitalizations of rural medicare beneficiaries. Health Aff. (Millwood). 38, 1993–2002. https://doi.org/10.1377/hltha ff.2019.00838 (2019).
6. OECD.
8. Yu, Q. et al. An MRI-based strategy for differentiation of frontotemporal dementia and Alzheimer's disease. *Alzheimers Res. Ther.* **13**, 23. https://doi.org/10.1186/s13195-020-00757-5 (2021).
10. Suk, H. I., Lee, S. W., Shen, D., Alzheimer's Disease Neuroimaging, I. Deep sparse multi-task learning for feature selection in multiclass diagnosis of Alzheimer's disease. *IEEE Trans. Biomed. Eng.* **62**, 1132–1140. https://doi.org/10.1109/TBME.2014.2372011 (2015).
12. Suk, H. I. & Shen, D. Alzheimer’s Disease Neuroimaging. I. Deep sparse multi-task learning for feature selection in Alzheimer's disease diagnosis. *Brain Struct. Funct.* **221**, 2569–2587. https://doi.org/10.1007/s00429-015-1059-y (2016).
14. Xiao, L. et al. Development of the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Group on Mild Cognitive Impairment. *Alzheimers Dement.* **7**, 263–269. https://doi.org/10.1016/j.jalz.2011.03.005 (2011).
15. Xiao, L. et al. Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. *Sci. Rep.* **8**, 3697. https://doi.org/10.1038/s41598-018-22871-z (2018).
17. Li, F., Liu, M. & Li, J. Automatic alzheimer's disease recognition from MRI data using deep learning method. *J. Appl. Math. Phys.* **5**, 7. https://doi.org/10.4236/jamp.2017.59159 (2017).
19. Xiao, L. & Shen, D. Landmark-based deep multi-instance learning for brain disease diagnosis. *Med. Image Anal.* **43**, 157–168. https://doi.org/10.1016/j.media.2017.10.005 (2018).
21. Wu, X., Yang, Z., Peng, J. & Zhou, J. Global denoising for 3D MRI. *IEEE Trans. Biomed. Eng.* **66**, 2858–2869. https://doi.org/10.1109/TBME.2019.2892434 (2019).
23. Wu, X., Yang, Z., Peng, J. & Zhou, J. Global denoising for 3D MRI. *Med. Image Anal.* **45**, 378–389. https://doi.org/10.1016/j.media.2019.07.009 (2019).
25. Smith, S. M. BET: Brain extraction tool. *Neuroimage Clin.* **3**, 101–118. https://doi.org/10.1016/j.nicl.2014.04.005 (2014).
27. Barthel, H., et al. Current status of brain MRI in the differentiation of dementia due to Alzheimer's disease and other dementias. *Brain* **139**, 1920–1933. https://doi.org/10.1093/brain/awaa137 (2020).
29. Jack, C. R. Jr. et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol.* **9**, 119–128. https://doi.org/10.1016/S1474-4422(09)70299-6 (2010).
31. Basso, A. et al. Alzheimer's disease: new perspectives for diagnosis and treatment. *Lancet Neurol.* **7**, 54–62. https://doi.org/10.1016/S1474-4422(06)70077-1 (2008).
33. Sperling, R. A., 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.* **7**, 263–269. https://doi.org/10.1016/j.jalz.2011.03.005 (2011).
35. McKhann, G. M. 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.* **7**, 270–279. https://doi.org/10.1016/j.jalz.2011.03.008 (2011).
37. Chauveau, L. et al. Medial temporal lobe subregional atrophy in aging and Alzheimer’s disease: A longitudinal study. *Front. Aging Neurosci.* **13**, 750154. https://doi.org/10.3389/fnagi.2021.750154 (2021).
39. Apostolova, L. G. et al. Conversion of mild cognitive impairment to Alzheimer disease predicted by hippocampal atrophy maps. *Arch. Neurol.* **63**, 693–699. https://doi.org/10.1001/archneur.63.5.693 (2006).
41. Lee, S. et al. Alzheimer's Disease Neuroimaging. I. Magnetic resonance imaging texture predicts progression to dementia due to Alzheimer disease earlier than hippocampal volume loss. *J. Psychiatry Neurosci.* **45**, 7–14. https://doi.org/10.1503/1801071 (2020).
43. Lin, C. K., et al. Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. *Sci. Rep.* **8**, 3697. https://doi.org/10.1038/s41598-018-22871-z (2018).
45. Jack, C. R. Jr. et al. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. *Neurology* **55**, 484–489. https://doi.org/10.1212/wnl.55.4.484 (2000).
47. Jack, C. R. Jr. et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol.* **9**, 119–128. https://doi.org/10.1016/S1474-4422(09)70299-6 (2010).
49. Lee, S., et al. Alzheimer's Disease Neuroimaging. I. Magnetic resonance imaging texture predicts progression to dementia due to Alzheimer disease earlier than hippocampal volume loss. *J. Psychiatry Neurosci.* **45**, 7–14. https://doi.org/10.1503/1801071 (2020).
45. Haussmann, R. et al. Precuneus structure changes in amnestic mild cognitive impairment. Am. J. Alzheimers Dis. Other Demen. 32, 22–26. https://doi.org/10.1177/1533317516678087 (2017).

46. Minoshima, S. et al. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. Ann. Neurol. 42, 85–94. https://doi.org/10.1002/ana.410420114 (1997).

47. Chetelat, G., Villain, N., Desgranges, B., Eustache, F. & Baron, J. C. Posterior cingulate hypometabolism in early Alzheimer's disease: What is the contribution of local atrophy versus disconnection?. Brain 132, e133. https://doi.org/10.1093/brain/awp253 (2009) (author reply e134).

Acknowledgements
The authors thank all members of the staff and participants at the Seoul National University Bundang Hospital, Seoul Asan Medical Center, and Korea Veteran Health Service for their important contributions.

Author contributions
J.S.K. and J.W.H. contributed conceptualization, methodology, formal analysis, investigation, resources, data curation and writing-original draft; J.B.B. contributed methodology, formal analysis and data curation; D.G.M., J.S., J.E.K., H.L., H.W.Y., E.L., J.Y.K., L.S., S.J.C., I.K., S.W.H., M.J.K., C.H.S., W.H.S. and S.J.K. contributed investigation and resources; D.L. contributed software; K.W.K. contributed conceptualization, methodology, formal analysis, writing-original draft and supervision.

Funding
This study was partly supported by Institute for Information and communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (2018-2-00861, Intelligent SW Technology Development for Medical Data Analysis) and by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) and Korea Dementia Research Center (KDRC), funded by the Ministry of Health & Welfare and Ministry of Science and ICT, Republic of Korea (HU20C0015).

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to K.W.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022