Received: 14 August 2024 | Revised: 3 October 2024 | Accepted: 17 October 2024

DOI: 10.1002/dad2.70037

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Artificial intelligence–based rapid brain volumetry substantially improves differential diagnosis in dementia

Jan Rudolph[1] | Johannes Rueckel[1,2] | Jörg Döpfert[3] | Wen Xin Ling[3] | Jens Opalka[4] | Christian Brem[2] | Nina Hesse[1] | Maria Ingenerf[1] | Vanessa Koliogiannis[1] | Olga Solyanik[1] | Boj F. Hoppe[1] | Hanna Zimmermann[2] | Wilhelm Flatz[1] | Robert Forbrig[2] | Maximilian Patzig[2] | Boris-Stephan Rauchmann[2,5,6] | Robert Perneczky[6,7] | Oliver Peters[8,9] | Josef Priller[8,10,11,12,13] | Anja Schneider[14,15] | Klaus Fliessbach[14,15] | Andreas Hermann[16,17] | Jens Wiltfang[18,19,20] | Frank Jessen[11,21,22] | Emrah Düzel[23,24] | Katharina Buerger[6,25] | Stefan Teipel[26,27] | Christoph Laske[28,29] | Matthis Synofzik[28,30] | Annika Spottke[14,31] | Michael Ewers[6] | Peter Dechent[32] | John-Dylan Haynes[33] | Johannes Levin[6,34,35] | Thomas Liebig[2] | Jens Ricke[1] | Michael Ingrisch[1,36] | Sophia Stoecklein[1]

Correspondence
Jan Rudolph, Department of Radiology, University Hospital LMU Munich, Marchioninistr. 15, 81377 Munich, Germany.
Email: jan.rudolph@med.uni-muenchen.de

Abstract

Introduction: This study evaluates the clinical value of a deep learning–based artificial intelligence (AI) system that performs rapid brain volumetry with automatic lobe segmentation and age- and sex-adjusted percentile comparisons.

Methods: Fifty-five patients—17 with Alzheimer's disease (AD), 18 with frontotemporal dementia (FTD), and 20 healthy controls—underwent cranial magnetic resonance imaging scans. Two board-certified neuroradiologists (BCNR), two board-certified radiologists (BCR), and three radiology residents (RR) assessed the scans twice: first without AI support and then with AI assistance.

Results: AI significantly improved diagnostic accuracy for AD (area under the curve −AI: 0.800, +AI: 0.926, $p < 0.05$), with increased correct diagnoses ($p < 0.01$) and reduced errors ($p < 0.03$). BCR and RR showed notable performance gains (BCR: $p < 0.04$; RR: $p < 0.02$). For the diagnosis FTD, overall consensus ($p < 0.01$), BCNR ($p < 0.02$), and BCR ($p < 0.05$) recorded significantly more correct diagnoses.

Discussion: AI-assisted volumetry improves diagnostic performance in differentiating AD and FTD, benefiting all reader groups, including BCNR.

KEYWORDS
Alzheimer's disease, artificial intelligence, brain volumetry, clinical cohorts, frontotemporal dementia

Highlights

- Artificial intelligence (AI)-supported brain volumetry significantly improved the diagnostic accuracy for Alzheimer's disease (AD) and frontotemporal dementia (FTD), with notable performance gains across radiologists of varying expertise levels.
- The presented AI tool is readily clinically available and reduces brain volumetry processing time from 12 to 24 hours to under 5 minutes, with full integration into picture archiving and communication systems, streamlining the workflow and facilitating real-time clinical decision making.
- AI-supported rapid brain volumetry has the potential to improve early diagnosis and to improve patient management.

1 | INTRODUCTION

Cognitive decline caused by dementia not only leads to suffering for affected patients and their caregivers but also carries substantial socioeconomic cost.[1–3] As Western societies rapidly age, the incidence and prevalence of dementia increase accordingly, signifying the urgent need for improved diagnosis and early detection.[4–7] In recent years, artificial intelligence (AI) has been increasingly used in health care, including in the diagnosis of Alzheimer's disease (AD) and other dementias. AI models can already analyze a wide range of diagnostic markers, including neuroimaging results, but also other factors such as lifestyle data.[8,9] Brain volumetry, as a tool for quantifying the volume of brain structures by layer-by-layer segmentation, can substantially contribute to the differential diagnosis of dementia patients by revealing anatomical patterns of atrophy.[10–12] It can also facilitate monitoring disease progression and treatment response, which is of particular interest as novel biological therapies for dementia are currently emerging.[13–16]

Automated brain volumetry based on three-dimensional magnetic resonance imaging (MRI) has long been established[17–19] but has not yet been included in routine radiological reporting of MRI scans of dementia patients. The reason is at least threefold. First, conventional segmentation and volumetry are time consuming and require dedicated computational resources. Second, the volumetry results are mostly not integrated in the picture archiving and communication system (PACS) environment of the radiologist and are thus not readily available for clinical decision making. Third, standardized, age- and sex-adapted reference values for each segmented brain structure are currently lacking.[20]

Here we analyze the clinical value of an AI-based brain segmentation tool using deep learning methodology, which reduces computational time from ≈ 12 to 24 hours using conventional segmentation to under 5 minutes (considering an ordinary consumer grade personal computer with a graphics processing unit). In combination with full PACS integration and provided reference values, this brain volumetry tool considerably increases clinical applicability and enables the radiologist to include and interpret volumetry results in their report.

Our study hypothesis was that radiologists with varying levels of training could enhance their performance in the differential diagnosis of dementia by integrating the supplementary information provided by brain volumetry into their routine clinical practice. To test this hypothesis, we conducted a reading study involving highly specialized board-certified neuroradiologists (BCNRs), board-certified radiologists (BCRs), and radiology residents (RRs) at various stages of residency training. These participants evaluated magnetic resonance (MR) images for the presence of AD and frontotemporal dementia (FTD), both with and without the support of AI-based automated brain segmentation, and comparison of volumetry results to age- and sex-adapted percentiles.

2 | MATERIALS AND METHODS

2.1 | Ethics statement

Investigators only used fully anonymized data. The acquisition of MRI data for the associated studies was approved by the local ethics committees of the involved sites of the German Center for Neurodegenerative Diseases (DZNE). DZNE approved the use of anonymized data for this study. The use of the applied artificial intelligence software for study purposes was approved by the institutional legal department of the University Hospital of the Ludwig Maximilian University of Munich. The study was conducted in accordance with the Declaration of Helsinki.

RESEARCH IN CONTEXT

1. **Systematic review**: We reviewed the literature on brain volumetry, searching PubMed and Google Scholar, with a focus on its role in the diagnosis of Alzheimer's disease (AD) and frontotemporal dementia (FTD). Despite the proven utility of brain volumetry in identifying anatomical patterns of atrophy, its adoption in routine clinical practice remains limited due to factors such as time-consuming segmentation processes, unavailable reference standards for volumetry results, and lack of picture archiving and communication system (PACS) integration.
2. **Interpretation**: Our study demonstrates that artificial intelligence (AI)-supported brain segmentation combined with the comparison of volumetry results to age- and sex-adapted reference values significantly enhances the diagnostic accuracy of radiologists, including those with extensive experience, in the differential diagnosis of AD and FTD. The demonstrated AI tool, with its rapid data processing and full PACS integration, facilitates the incorporation of volumetric data into routine clinical reporting, which might lead to more accurate diagnoses and potentially earlier detection of dementia, which might ultimately improve patient management.
3. **Future directions**: Further research should focus on evaluating the long-term clinical benefits of AI-assisted brain volumetry, including its impact on patient outcomes and treatment pathways.

2.2 | Consent statement

All human study participants gave informed consent for the use of their clinical, diagnostic, and imaging data as part of DZNE-related studies.

2.3 | Patient cohort

Fully anonymized cranial MRI data of 161 participants were obtained from the DZNE. The datasets are part of the study cohorts DELCODE (DZNE—Longitudinal Cognitive Impairment and Dementia Study), DESCRIBE-FTD (DZNE—Clinical Registry Study on Frontotemporal Dementia), and DANCER (Degeneration Controls and Relatives). Images were acquired at different centers of the DZNE, resulting in heterogeneity of MRI scanners. All involved investigators in our data analysis had no access to any personal information of the patients (name, birthday, address, etc.).

In all delivered cases, clinically confirmed diagnosis of either AD, FTD, or no neurodegenerative disease/healthy control was available. Patient characteristics, including clinical test results, are presented in Table 1. Among the FTD cases, the following variants are represented: behavioral variant of frontotemporal dementia (bvFTD), progressive non-fluent aphasia (PNFA), and semantic dementia (SemD). The diagnosis of possible AD was based on the diagnostic criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA).[21] Disease assessment included a standardized medical history including medications taken; a comprehensive clinical neurological examination including various established tests of memory, language, and motor function; a systematic analysis of biomaterials (including recurrent blood, urine, and cerebrospinal fluid [CSF] diagnostics); biomarker diagnostics (including amyloid beta [Aβ]42, Aβ40, tau, phosphorylated tau [p-tau]); and imaging diagnostics such as MRI and nuclear medicine examinations (e.g., amyloid positron emission tomography). For detailed information on data analysis in the DELCODE cohort, see Jessen et al.[22]

In the FTD cohort, Rascovsky et al.'s diagnosis criteria[23] were used for the diagnosis of bvFTD, and Gorno-Tempini et al.'s criteria[24] were used for primary progressive aphasias (PNFA and SemD). The spectrum of investigations in the FTD cohort included clinical examination, medical history, concomitant medication, as well as biomarker diagnostics (e.g., Aβ42, Aβ40, tau, p-tau in CSF), genetics (whole genome/whole exome, C9orf, progranulin, tau), imaging diagnostics such as MRI, and a large battery of neuropsychological tests, some of which overlap with the DELCODE cohort.

Further image preselection for this study was conducted by an experienced radiology resident with 5 years of experience in neuroradiology. Inclusion criteria were: presence of a three-dimensional T1-weighted sequence (3D T1w) and a 3D fluid-attenuated inversion recovery sequence (FLAIR) acquired in native technology (without prior injection of contrast agent), sufficient image quality (visual check by the radiology resident, e.g., no motion or susceptibility artifacts), no white matter lesions $\geq$ Fazekas grade 2, no sign of prior ischemic infarction, and successful completion of quality check by AI tool according to Digital Imaging and Communications in Medicine tag–based image criteria [see Table S1 in supporting information]). The resulting collective included 17 patients with AD, 18 patients with FTD (12 bvFTD, 5 PFNA, and 1 SemD), and 20 healthy controls. All images were acquired on 3 T scanners. For MR imaging parameters, please refer to Table S2 in supporting information.

2.4 | Reading/questionnaire/consensus formation

3D T1w and 3D FLAIR sequences were case-wise extracted and handed over to seven readers—two BCNRs (BCNR I: 10 years of experience in image level dementia differential diagnostics (YOE) and BCNR II: 10 YOE), two BCR (BCR I: 7 YOE and BCR II: 5 YOE), and three RR (RR; RR I: 0 YOE, RR II: 1 YOE, and RR III: 1 YOE, including a special neuroradiologic training). All readers were blinded to the individual's diagnosis. Additionally, 3D T1w images were sent to the fully PACS integrated pipeline of the AI system mdbrain (Mediaire GmbH, version 3.3.0), installed at our institution (see section 2.6). The AI tool functioned uniformly across all scans, returning user readable results.

TABLE 1 Clinical parameter of enrolled patients sorted by study groups.

Parameter / cohorts	Alzheimer's disease	Frontotemporal dementia	Healthy controls
n / %	17 / 30.9%	18 / 32.7%	20 / 36.4%
Age at examination (mean / standard deviation / range)	74.5 / 5.9 / 64–87 years	63.2 / 9.9 / 41–80 years	68.6 / 7.6 / 55–80
Sex (female count / %)	13 / 76.5%	5 / 27.8%	11 / 55.5%
Body mass index (mean / standard deviation / range)	25.0 / 3.4 / 21–32 kg/m^2	25.1 / 3.3 / 19–30 kg/m^2	26.3 / 2.9 / 22–30 kg/m^2
Smoker (n / % / mean pack years of smoker)	0 / 0% / none	5 / 27.8% / 22.3 pack years	1 / 5.0% / 2 pack years
Educational years (mean / standard deviation / range)	12.5 / 1.9 / 10–18 years	14.4 / 2.5 / 12–19 years	14.4 / 2.7 / 11–19 years
Scores in clinical tests (mean / standard deviation / range)			
Mini Mental State Examination (MMSE)	23.2 / 3.5 / 16–28	22.7 / 6.8 / 5–29	29.2 / 0.8 / 28–30
Geriatric Depression Scale (GDS)	1.3 / 1.6 / 0–5	3.4 / 4.1 / 0–14	0.6 / 0.6 / 0–2
Neuropsychiatric Inventory Questionnaire (NPI-Q)	3.4 / 3.4 / 0–12	11.2 / 7.4 / 2–23	0.4 / 0.8 / 0–2
Clinical Dementia Rating Global (CDR)	0.8 / 0.4 / 0.5–2	1.0 / 0.7 / 0–2	0.0 / 0.1 / 0–0.5
Clinical Dementia Rating Sum of Boxes (CDR-SB)	4.4 / 2.8 / 0.5–12	6.5 / 4.1 / 0.5–14	0.1 / 0.2 / 0–0.5

Results were displayed in the form of individual "volumetry reports" (for a truncated version, see Figure 1). The reports were exported as PDF files and handed over to the readers during the reading process. All readers had to evaluate the cases in a two-step reading procedure. Cases were assessed in terms of the suspected pathology present using a likelihood score (0–5), adding up to a sum of 5 for all three diagnostic categories (AD, FTD, and healthy control). Here are two examples: If the reader is certain that FTD is present, he/she may award 5 points to FTD but must award 0 for the AD and healthy control categories. If the reader is unsure whether AD or FTD is present, he/she can, for example, assign 2 points to AD and FTD, respectively, and 1 point to healthy control. A low, odd number that cannot be divided by 3 was deliberately chosen because this requires a stronger weighting of the differential diagnoses. Reading I was a diagnostic evaluation of the bare 3D T1w and 3D FLAIR images without any AI support. In Reading II, all readers had to evaluate the images again, given the additional information of the AI/volumetry reports. To reduce confirmation bias, readers received the volumetry reports after finishing Reading I, had to hand back evaluation sheets from Reading I prior to Reading II, were given a wash-out phase of 30 to 60 days, and were blinded for Reading I likelihood scores during the ongoing Reading II. After Reading II, readers completed a questionnaire asking for their individual assessment of the AI tool. For subgroup analysis, reader consensus was formed by summing up the individual likelihood scores of the readers in the specific subgroup: overall consensus (OC) considers all seven participating readers, BCNR consensus considers the two BCNRs, BCR consensus considers the two BCRs, and RR consensus considers the three RRs.

2.5 | Statistical analysis

Individual performance and reader consensus performance (based on the likelihood scores) with and without AI support were statistically evaluated using receiver operating characteristic (ROC) curve with calculation of the area under the ROC curve (AUC) and operating point optimization with the Youden *J* statistic. This statistical methodology has already been established in several studies.[25–29] ROC curves were compared with paired ROC tests. Comparisons of assigned diagnosis points according to likelihood scores were performed using a Student *t* test for paired samples. Questionnaires with Likert-scaled choices were displayed using histograms. All statistical calculations and graphic illustrations were performed using open-source programming language R.[30] The R code can be made available upon request.

2.6 | AI algorithm

For quantitative analysis of brain MRIs, the commercially available AI tool "mdbrain," version 3.3.0 (Mediaire GmbH) was used. The system leverages a custom deep learning segmentation model based on the U-Net architecture[31] to perform a highly accurate side- and region-specific (e.g., lobes) rapid brain volumetry, which was trained on a heterogeneous dataset of 3D T1w images (n = 1851, balanced male/female). Augmentation techniques (augmentation of contrast, resolution, rotation, and elastic deformation) have been used to maximize the model's applicability in daily routine.

The volumes of 18 brain regions, including the hippocampus, are determined and percentiles are derived by comparison to a cohort of healthy individuals (n = 3179, balanced male/female, age range 18–92), while accounting for age, sex, and total intracranial volume. Volumes and percentiles are displayed in tabular format, along with clinically relevant MRI slices and schematics to highlight pathological values beyond two standard deviations from the mean (example report provided in Figure 1).

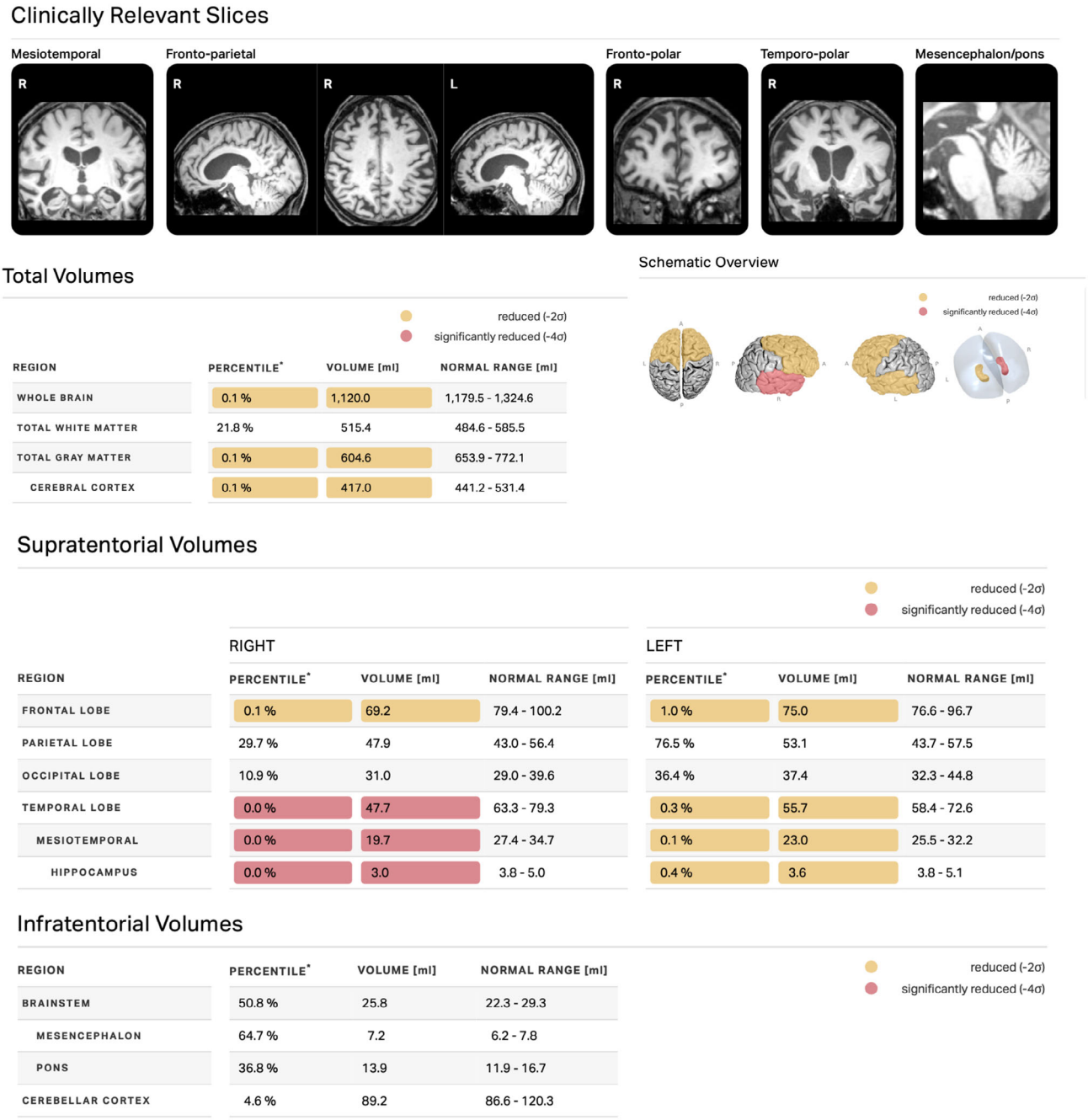

FIGURE 1 Artificial intelligence–based volumetry report. Excerpt from a volumetry report of a study patient diagnosed with frontotemporal dementia. The measured volumes from in total 18 brain regions (including the hippocampus) are determined from the segmentation and then automatically compared against an age- and sex-matched reference cohort of healthy individuals ($n = 3179$) to yield percentiles. The entire analysis runs in < 5 minutes. Volumes and percentiles are displayed in tabular format, along with clinically relevant magnetic resonance imaging slices and schematics to highlight pathological values beyond two standard deviations from the mean.

3 | RESULTS

3.1 | ROC analysis and distribution of diagnostic points

ROC analysis and comparison tests of correctly assigned diagnostic points showed that correct decision making was substantially improved by AI support. ROC curves with according to Youden *J* statistics optimized operation points and further metrics at the optimized operation point (acc [accuracy], sens [sensitivity], spec [specificity], ppv [positive predictive value], npv [negative predictive value]), fpr [false positive rate], fnr [false negative rate], and ctl [closest top left]) are depicted in Figures 2 (AD) and 3 (FTD), with part A of both figures showing subfigures without AI support and part B showing subfigures with AI support. Comparisons of AUCs and correctly/incorrectly assigned diagnostic points can be found in Tables 2 (AD) and 3 (FTD).

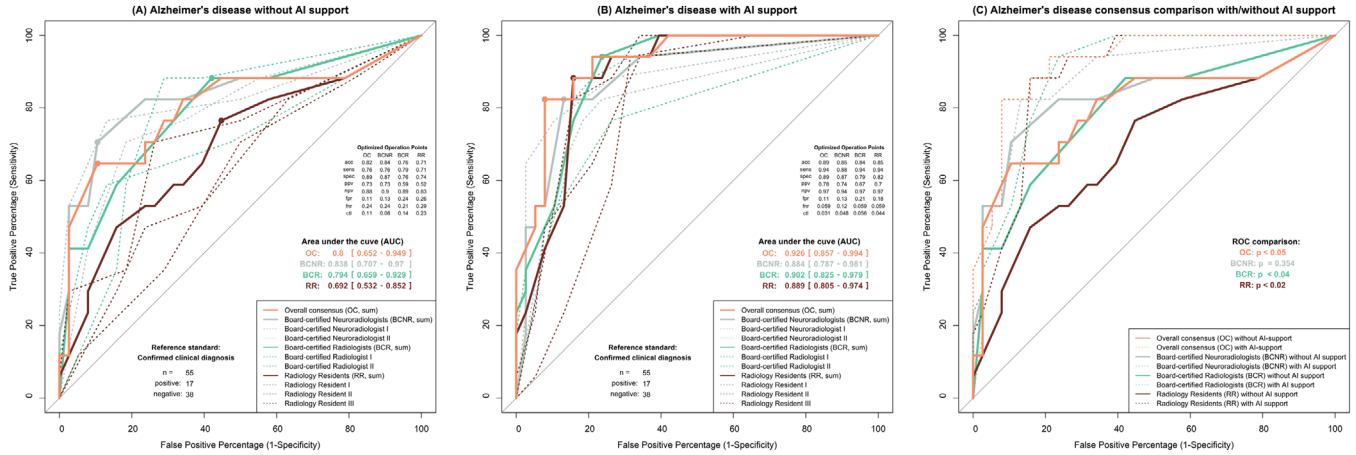

FIGURE 2 Performance in the radiological differential diagnosis of AD without (A) and with AI support (B) and performance comparison of consensus (C). OC, BCR, and RR consensus improved significantly with AI support, with OC reaching an AUC of up to 0.926. BCNR consensus improved slightly (not statistically significantly) with AI support. acc, accuracy; AI, artificial intelligence; AD, Alzheimer's disease; AUC, area under receiver operating characteristic curve; BCR, board-certified radiologist; BCNR, board-certified neuroradiologist; ctl, closest top left; fnr, false negative rate; fpr, false positive rate; npv, negative predictive value; OC, overall consensus; ppv, positive predictive value; RR, radiology resident; sens, sensitivity; spec, specificity.

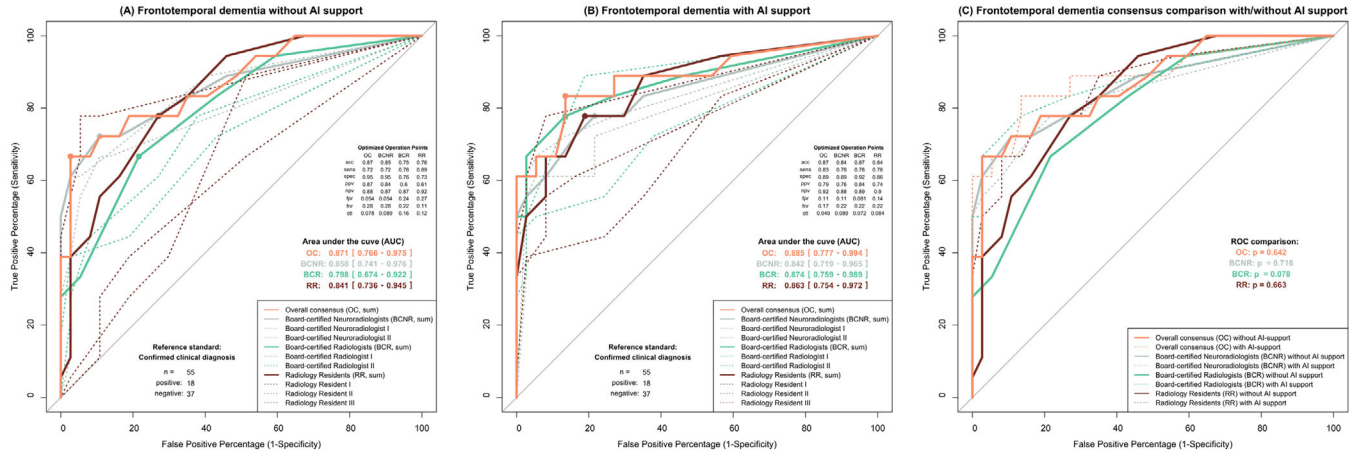

FIGURE 3 Performance in the radiological differential diagnosis of FTD without (A) and with AI support (B) and performance comparison of consensus (C). Consensus improved for the most part with AI support (not statistically significant). Outliers with poor performance become fewer with AI support. acc, accuracy; AI, artificial intelligence; AUC, area under receiver operating characteristic curve; BCR, board-certified radiologist; BCNR, board-certified neuroradiologist; ctl, closest top left; fnr, false negative rate; fpr, false positive rate; FTD, frontotemporal dementia; npv, negative predictive value; OC, overall consensus; ppv, positive predictive value; RR, radiology resident; sens, sensitivity; spec, specificity.

Evaluations for the correct detection of healthy controls can be found in Figure S1 and Table S3 in supporting information.

AD diagnosis was clearly improved by AI support, showing a statistically significant performance improvement in OC ($p < 0.05$) with a resulting AUC with AI support (+AI) of 0.926 compared to an AUC of 0.800 without AI support (−AI) and an (according to Youden *J* statistic optimized) accuracy of 0.89 with AI support compared to an accuracy of 0.82 without AI support. Considering a potential maximum of 275 points (likelihood score 0–5; 5 points/case × 55 cases) and a maximum number of correct points of 85 for AD diagnosis (17 correct cases), overall reader consensus allocated 10.5 additional points correctly to the AD diagnosis. This increase in correctly distributed points was statistically significant ($p < 0.01$). Also, reduction of incorrectly assigned AD diagnostic points was statistically significant (−AI: 42.1, +AI: 32.1, $p < 0.03$). In the subgroup analysis, performance improvement was statistically significant for BCR (AUCs: −AI: 0.794, +AI: 0.902, $p < 0.04$) and RR consensus (AUCs: −AI: 0.692, +AI: 0.889, $p < 0.02$), while in BCNR consensus performance non-significantly exceeded the results without AI support (AUCs −AI: 0.838, +AI: 0.884, $p = 0.354$). Looking at the assigned points, RR consensus assigned 17.3/48% more points to the correct AD diagnosis (−AI: 36.0, + AI: 53.3, $p < 0.01$) and non-significantly decreased the number of incorrectly assigned diagnostic

TABLE 2 Individual and consensus performance in the detection of Alzheimer's disease. AUCs are given with 95% CI in parentheses.

	AUC −AI	AUC +AI	*p*	Correct points −AI	Correct points +AI	*p*	False points −AI	False points +AI	*p*
Overall consensus	0.800 (0.652–0.949)	0.926 (0.857–0.994)	**<0.05**	40.6	51.1	**<0.01**	42.1	32.1	**<0.03**
BCNR consensus	0.838 (0.707–0.970)	0.884 (0.787–0.981)	0.354	47.0	53.0	0.159	36.0	22.5	**<0.01**
BCNR I	0.786 (0.650–0.922)	0.822 (0.701–0.943)	0.303	54	55	0.859	47	28	**<0.01**
BCNR II	0.827 (0.686–0.967)	0.882 (0.776–0.989)	0.297	40	51	**<0.02**	25	17	0.088
BCR consensus	0.794 (0.659–0.929)	0.902 (0.825–0.979)	**<0.04**	41.0	46.0	0.151	34.5	24.0	**<0.04**
BCR I	0.782 (0.652–0.909)	0.866 (0.765–0.967)	0.053	50	62	0.070	43	34	0.228
BCR II	0.721 (0.562–0.881)	0.793 (0.662–0.923)	0.134	32	30	0.622	26	14	0.101
RR consensus	0.692 (0.532–0.852)	0.889 (0.805–0.974)	**<0.02**	36.0	53.3	**<0.01**	51.3	44.0	0.350
RR I	0.594 (0.435–0.754)	0.782 (0.664–0.900)	**<0.04**	35	48	**<0.05**	63	54	0.355
RR II	0.629 (0.471–0.786)	0.885 (0.802–0.969)	**<0.01**	36	72	**<0.001**	54	52	0.869
RR III	0.713 (0.560–0.866)	0.834 (0.725–0.942)	0.090	37	40	0.666	37	26	0.188

Note: Because $n = 17$ and 5 points can be assigned per case, the maximum number of correct points was 85. The total number of points to assign was 275. The number of incorrect points refers to the total number of AD points incorrectly assigned in FTD cases or in healthy controls. Correct and incorrectly assigned points therefore do not necessarily add up to 85. *P* values are calculated with DeLong test for correlated ROC curves and paired Student *t* test. Statistically significant *p* values ($p < 0.05$) are marked in bold print. −AI: Results without AI support, +AI: Results with AI support.

Abbreviations: AI, artificial intelligence; +AI, with AI support; −AI, without AI support; AD, Alzheimer's disease; AUC, area under receiver operating characteristic curve; BCR, board-certified radiologist; BCNR, board-certified neuroradiologist; CI, confidence interval; FTD, frontotemporal dementia; ROC, receiver operating characteristic; RR, radiology resident.

points (−AI: 51.3, +AI: 44.0, $p = 0.350$). In BCR consensus correctly assigned points tended to increase with AI support (−AI: 41.0, +AI: 46.0, $p = 0.151$), and incorrectly assigned points decreased significantly (*-AI: 34.5, +AI: 24.0, P* < 0.04). In BCNR consensus, AD diagnostic points falsely assigned to non-AD cases were reduced significantly with AI support (-AI: 36.0, +AI: 22.5, $p < 0.01$). In the evaluation of individual readers, the most pronounced effects can be observed in the most unexperienced readers, RR I and RR II, who have statistically significantly improved their performance (AUCs; RR I: −AI: 0.594, + AI: 0.782, $p < 0.04$; RR II: −AI: 0.629, +AI: 0.885, $p < 0.01$) and increased the number of correctly assigned AD diagnostic points (RR I: −AI: 35, +AI: 48, $p < 0.05$; RR II:, −AI: 36, +AI: 72, $p < 0.001$), whereas RR II doubled the correctly assigned points. RR III as a more experienced radiology resident with special neuroradiological training showed non-significant performance increase (AUCs, −AI: 0.713, +AI: 0.834, $p = 0.090$), which is comparable to the individual BCNR and BCR readers. Notably, even the highly experienced BCNR I significantly reduced the number of falsely assigned diagnostic points (−AI: 47, +AI: 28, $p < 0.01$) and the highly experienced BCNR II significantly improved the number of correctly assigned points (−AI: 40, +AI: 51, $p < 0.02$).

In FTD diagnosis the overall consensus slightly improved with AI support (AUCs, −AI: 0.871, +AI: 0.885, $p = 0.642$) and correctly assigned points were significantly higher (−AI: 39.0, +AI: 46.1, $p < 0.01$). The accuracy (optimized according to Youden *J* statistic) was comparable with and without AI support (−AI: 0.87, +AI: 0.87). Incorrectly assigned points tended to be lower (−AI: 23.6, +AI: 19.9, $p = 0.348$). For the BCNR and BCR consensus, although AUC differences were not statistically significant (AUCs; BCNR: −AI: 0.858, +AI: 0.842, $p = 0.716$; BCR: −AI: 0.789, +AI: 0.874, $p = 0.078$), correctly assigned points were significantly higher with AI support (BCNR: −AI: 46.0, +AI: 56.0, $p < 0.02$; BCR: −AI: 36.5, +AI: 42.0, $p < 0.05$). Incorrectly assigned points showed no statistically different results. In RR consensus AUC and points comparisons showed no statistically significant differences. Considering the individual readers, BCNR I and BCNR II significantly improved the number of correctly assigned FTD diagnostic points (BCNR I: −AI: 46, +AI: 56, $p < 0.04$; BCNR II: −AI: 46, +AI: 56, $p < 0.04$). BCR II showed a statistically significantly higher performance with AI support (AUCs, −AI: 0.706, +AI: 0.887, $p < 0.01$) and a significant reduction of incorrectly assigned diagnostic points (−AI: 24, +AI: 9, $p < 0.02$). Also, RR II could significantly

TABLE 3 Individual and consensus performance in the detection of Frontotemporal dementia. AUCs are given with 95% CI in parentheses.

	AUC −AI	AUC +AI	*p*	Correct points −AI	Correct points +AI	*p*	False points −AI	False points +AI	*p*
Overall consensus	0.871 (0.766–0.975)	0.885 (0.777–0.994)	0.642	39.0	46.1	**<0.01**	23.6	19.9	0.348
BCNR consensus	0.858 (0.740–0.976)	0.842 (0.719–0.965)	0.716	46.0	56.0	**<0.02**	16.0	22.0	0.165
BCNR I	0.794 (0.666–0.921)	0.797 (0.667–0.928)	0.894	46	56	**<0.04**	12	25	0.052
BCNR II	0.843 (0.728–0.958)	0.828 (0.701–0.955)	0.792	46	56	**<0.04**	20	19	0.830
BCR consensus	0.798 (0.674–0.922)	0.874 (0.759–0.989)	0.078	36.5	42.0	**<0.05**	26.5	17.0	0.076
BCR I	0.750 (0.611–0.889)	0.747 (0.600–0.894)	0.957	42	47	0.358	29	25	0.632
BCR II	0.706 (0.555–0.856)	0.887 (0.792–0.981)	**<0.01**	31	37	0.057	24	9	**<0.02**
RR consensus	0.841 (0.736–0.945)	0.863 (0.754–0.972)	0.663	36.0	42.3	0.187	26.7	20.3	0.239
RR I	0.661 (0.520–0.803)	0.697 (0.544–0.849)	0.703	29	41	0.135	38	44	0.513
RR II	0.610 (0.453–0.766)	0.754 (0.620–0.888)	0.074	26	37	0.154	34	12	**<0.03**
RR III	0.865 (0.755–0.975)	0.868 (0.761–0.975)	0.943	53	49	0.532	8	5	0.261

Note: Because $n = 18$ and 5 points can be assigned per case, the maximum number of correct points was 90. The total number of points to assign was 275. The number of incorrect points refers to the total number of FTD points incorrectly assigned in AD cases or in healthy controls. Correct and incorrectly assigned points therefore do not necessarily add up to 90. *P* values are calculated with DeLong test for correlated ROC curves and paired Student *t* test. Statistically significant *p* values ($p < 0.05$) are marked in bold print. −AI: results without AI support, +AI: results with AI support.

Abbreviations: AI, artificial intelligence; +AI, with AI support; −AI, without AI support; AD, Alzheimer's disease; AUC, area under receiver operating characteristic curve; BCR, board-certified radiologist; BCNR, board-certified neuroradiologist; CI, confidence interval; FTD, frontotemporal dementia; ROC, receiver operating characteristic; RR, radiology resident.

reduce the number of false FTD diagnostic points (–AI: 34, +AI: 12, $p < 0.03$).

3.2 | Reader's Questionnaire

Details on the Reader's Questionnaire and responses are provided in Figure S2 and Text S1 in supporting information.

4 | DISCUSSION

In this study, we found that all participating radiology readers improved their performance in diagnosing AD using AI-assisted rapid brain volumetry. The greatest effect was seen in BCRs and RRs. In the diagnosis of FTD, we found an increase in diagnostic confidence with AI support. Subjectively, readers also felt that their diagnostic confidence had improved using the AI-assisted rapid brain volumetry, as noted in a follow-up survey.

Specifically, we found that OC assessment of all readers showed a statistically significant improvement in performance with AI support for the diagnosis of AD, with a significant increase in the number of correctly assigned diagnostic points and a significant decrease in incorrectly assigned points. Looking at the individual subgroups, it became apparent that the performance with AI support for the differential diagnosis of AD increased significantly in the BCR and RR readers alike, with corresponding changes in the distribution of diagnostic points. In the RR readers, there were differences between two very inexperienced readers and one reader previously specifically trained in neuroradiology, who already showed a very good performance without AI support. The highly skilled BCNR also performed very well without AI support but still tended to improve slightly with AI support (not statistically significant).

For the differential diagnosis of FTD, although we did not find a statistically significant improvement in AUC with AI support, we did find a significant increase in the number of correctly assigned diagnostic points. Specifically, the increase could be noted in the OC and the group of highly skilled BCNRs. Some individual readers, moreover, were able to reduce the number of incorrect assignments.

Furthermore, without AI support, AUC values were generally higher in the FTD cohort than in the AD cohort. Therefore, it can be assumed that most readers more easily recognized the

frontotemporal changes characteristic of FTD than the temporal, mesiotemporal, or hippocampal changes typical for AD.

The improvement in performance and safety with the help of the AI tool is also accompanied by an improvement in reporting quality. With more precise findings, an earlier decision in diagnostics can promote potential therapy options and thus improve patient care. The presented AI solution is readily available and offers full PACS integration, which allows for a seamless workflow during routine examinations. With a short time of < 5 minutes for the common volumetric analysis, radiologists would be able to add the data to almost every cranial MRI report, thereby increasing sensitivity for even subtle volumetric changes and assumably facilitating early, potentially even preclinical diagnosis. Early diagnosis is essential, especially in view of the development of many new and evolving targeted therapies. For example, a recent phase III study with the compound lecanemab showed a reduction in amyloid markers in early AD and resulted in moderately less deterioration of cognitive abilities and functions after 18 months compared to the placebo group;[13] the drug has now been approved by the US Food and Drug Administration. With a steadily increasing workload due to more examinations and an equally non-increasing number of radiologists,[32] manual brain volumetry as a time-consuming methodology would not be possible in routine clinical practice for capacity reasons. Thus, the presented AI tool drastically increases the information content and diagnostic accuracy without compromising the valuable working time of radiologists.

An important asset of the present study is the diversity of different radiology readers from the neuroradiology and radiology department with on the one hand, inexperienced novices (0 years) and on the other hand highly experienced BCNRs (10 years), with other intermediate training levels being represented as well. In addition, the evaluation was based on a small, but clinically very well validated, study population. Patient data were collected according to a standardized multicenter procedure that has been established in recent studies.[22,33–38] The sample is clinically representative with regard to variance of age. The reference standard was very precisely defined on a clinical level and not based solely on image findings, as is common in the majority of other radiological reading studies for external validation of AI algorithms.[39,40] With AD and FTD, two very common types of dementia have been evaluated. With an incidence of 0.0 to 0.3 per 1000 person-years and ≈ 3% of all dementia cases, FTD is not a very common disease, but it is a common form of dementia in those < 65 years of age, accounting for ≈ 10%.[41] Methodologically, several techniques were used to reduce the risk of confirmation bias—for example, each reader had to hand in his/her results from Reading I and was allowed to start Reading II only after a wash-out phase. With 3D T1w and 3D FLAIR sequences, the readers had extensive data available without AI support, which in this form already allowed for morphological differential diagnosis.

Limitations of the study include the low total number of cases, a residual risk of confirmation bias (despite methodological consideration), a monocentric reading (with two departments: neuroradiology and radiology) in which some training effect may be present (e.g., BCRs teach RRs), and a pre-selection of cases that might favor less ambiguous cases. Another limitation captures the interpolation of ROC curves by the choice of a scaled scoring system. AUC values are therefore approximated. The radiological reading process without clinical information does not emulate the clinical workflow perfectly. The aim of our study was to evaluate AI-based rapid brain volumetry of MRI data in the assessment of dementia patients. Additional clinical information and interdisciplinary assessment of all available patient information should further improve diagnostic performance.

5 | CONCLUSION

With a plethora of AI solutions being offered to support the radiological workflow, critical assessment of added clinical value by radiologists and referring physicians is crucial. Here, we showed that AI-supported brain volumetry can add significant clinical value for the differential diagnosis of dementia. The majority of (even well-trained) radiologists can significantly increase their diagnostic accuracy in detecting AD and gain more confidence to commit to the diagnosis of FTD through the AI tool. This might lead to earlier diagnosis and, therefore, optimized patient management. The presented AI tool is readily available in the reporting workflow, offering rapid data processing and full PACS integration.

AUTHOR CONTRIBUTIONS

Jan Rudolph, Johannes Rueckel, Michael Ingrisch, and Sophia Stoecklein developed the study design. Jens Ricke, Thomas Liebig, and Sophia Stoecklein supervised the study. Robert Perneczky, Oliver Peters, Josef Priller, Anja Schneider, Klaus Fliessbach, Andreas Hermann, Jens Wiltfang, Frank Jessen, Emrah Düzel, Katharina Buerger, Stefan Teipel, Christoph Laske, Matthis Synofzik, Annika Spottke, Michael Ewers, Peter Dechent, John-Dylan Haynes, and Johannes Levin realized image acquisition as PIs of the individual DZNE sites. Johannes Rueckel, Boris-Stephan Rauchmann, and Nina Hesse applied for the DZNE data sets. Christian Brem preselected the image datasets. Jan Rudolph, Johannes Rueckel, Maria Ingenerf, Vanessa Koliogiannis, Hanna Zimmermann, Robert Forbrig, and Maximilian Patzig radiologically assessed the image data. Jörg Döpfert and Wen Xin Ling technically developed the presented AI solution. Jens Opalka medically supervised AI development. Michael Ingrisch assisted with the technical implementation of the AI solution. Jan Rudolph statistically analyzed and graphically illustrated the results. Johannes Rueckel and Boj F. Hoppe assisted with statistical analysis. Jan Rudolph wrote the initial manuscript draft. Sophia Stoecklein assisted with manuscript preparation. All co-authors critically reviewed the manuscript and have read and agreed to the submitted version of the manuscript.

AFFILIATIONS

[1]Department of Radiology, University Hospital, LMU Munich, Munich, Germany

[2]Department of Neuroradiology, University Hospital, LMU Munich, Munich, Germany

[3]Development Department, Mediaire GmbH, Berlin, Germany

[4]Medical Department, Mediaire GmbH, Berlin, Germany

[5]Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK

[6]German Center for Neurodegenerative Diseases (DZNE) Munich, Munich, Germany

[7]Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich, Germany

[8]German Center for Neurodegenerative Diseases (DZNE) Berlin, Berlin, Germany

[9]Department of Psychiatry, Campus Benjamin Franklin, Charité-Universitätsmedizin, Berlin, Germany

[10]Department of Psychiatry and Psychotherapy, Charité Berlin, Berlin, Germany

[11]Department of Psychiatry and Psychotherapy, Klinikum rechts der Isar, Technical University Munich, Munich, Germany

[12]UK Dementia Research Institute at the University of Edinburgh, Edinburgh Medical School, Edinburgh, UK

[13]Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK

[14]German Center for Neurodegenerative Diseases (DZNE) Bonn, Bonn, Germany

[15]Department of Old Age Psychiatry and Cognitive Disorders, University Hospital Bonn, University of Bonn, Bonn, Germany

[16]German Center for Neurodegenerative Diseases (DZNE) Rostock/Greifswald, Rostock, Germany

[17]Translational Neurodegeneration Section "Albrecht Kossel", Department of Neurology, Rostock University Medical Center, Rostock, Germany

[18]German Center for Neurodegenerative Diseases (DZNE) Goettingen, Goettingen, Germany

[19]Department of Psychiatry and Psychotherapy, University Medical Center Goettingen, University of Goettingen, Goettingen, Germany

[20]Neurosciences and Signaling Group, Institute of Biomedicine (iBiMED), Department of Medical Sciences, Campus Universitário de Santiago, University of Aveiro, Aveiro, Portugal

[21]Department of Psychiatry, Medical Faculty, University of Cologne, Cologne, Germany

[22]Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Cologne, Germany

[23]German Center for Neurodegenerative Diseases (DZNE) Magdeburg, Magdeburg, Germany

[24]Institute of Cognitive Neurology and Dementia Research (IKND), Otto-von-Guericke University Magdeburg, Magdeburg, Germany

[25]Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, Munich, Germany

[26]German Center for Neurodegenerative Diseases (DZNE) Rostock, Rostock, Germany

[27]Department of Psychosomatic Medicine, Rostock University Medical Center, Rostock, Germany

[28]German Center for Neurodegenerative Diseases (DZNE) Tuebingen, Tuebingen, Germany

[29]Section for Dementia Research, Hertie Institute for Clinical Brain Research and Department of Psychiatry and Psychotherapy, University of Tuebingen, Tuebingen, Germany

[30]Center of Neurology, Department of Neurodegenerative Diseases, University Hospital Tuebingen, Tuebingen, Germany

[31]Department of Neurology, University of Bonn, Bonn, Germany

[32]MR-Research in Neurosciences, Department of Cognitive Neurology, Georg-August-University Goettingen, Goettingen, Germany

[33]Bernstein Center for Computational Neuroscience, Charité-Universitätsmedizin, Berlin, Germany

[34]Department of Neurology, University Hospital, LMU Munich, Munich, Germany

[35]Munich Cluster of Systems Neurology (SyNergy), Munich, Germany

[36]Munich Center for Machine Learning (MCML), Munich, Germany

ACKNOWLEDGMENTS

Department of Radiology, University Hospital, LMU Munich, Germany received funding (discounted software usage) within a research cooperation from Mediaire GmbH.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

No individual disclosures to the content of this manuscript. Author disclosures are available in the supporting information.

ORCID

Jan Rudolph https://orcid.org/0000-0002-4849-8034

REFERENCES

1. Sopina E, Spackman E, Martikainen J, Waldemar G, Sørensen J. Long-term medical costs of Alzheimer's disease: matched cohort analysis. *Eur J Health Econ*. 2019;20:333-342. doi:10.1007/s10198-018-1004-0
2. Marešová P, Mohelská H, Dolejš J, Kuča K. Socio-economic Aspects of Alzheimer's Disease. *Curr Alzheimer Res*. 2015;12:903-911. doi:10.2174/1567205012091510191114 48
3. Morris ZA, Zaidi A, McGarity S. The extra costs associated with a cognitive impairment: estimates from 15 OECD countries. *Eur J Public Health*. 2021;31:647-652. doi:10.1093/eurpub/ckab011
4. Hou Y, Dan X, Babbar M, et al. Ageing as a risk factor for neurodegenerative disease. *Nat Rev Neurol*. 2019;15:565-581. doi:10.1038/s41582-019-0244-7
5. Feigin VL, Vos T, Nichols E, et al. The global burden of neurological disorders: translating evidence into policy. *Lancet Neurol*. 2020;19:255-265. doi:10.1016/S1474-4422(19)30411-9
6. Prince M, Wimo A, Guerchet M, et al. *World Alzheimer Report 2015*. Alzheimer's Disease International; 2015.
7. GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 2019;18:459-480. doi:10.1016/S1474-4422(18)30499-X
8. Newby D, Orgeta V, Marshall CR, et al. Artificial intelligence for dementia prevention. *Alzheimers Dement*. 2023;19:5952-5969. doi:10.1002/alz.13463
9. Akbarifar A, Maghsoudpour A, Mohammadian F, Mohammadzaheri M, Ghaemi O. Multimodal dementia identification using lifestyle and brain lesions, a machine learning approach. *AIP Advances*. 2024;14:065026. doi:10.1063/5.0211527
10. Giorgio A, De Stefano N. Clinical use of brain volumetry. *J Magn Reson Imaging*. 2013;37:1-14. doi:10.1002/jmri.23671
11. Pini L, Pievani M, Bocchetta M, et al. Brain atrophy in Alzheimer's Disease and aging. *Ageing Res Rev*. 2016;30:25-48. doi:10.1016/j.arr.2016.01.002
12. Velickaite V, Ferreira D, Lind L, et al. Visual rating versus volumetry of regional brain atrophy and longitudinal changes over a 5-year period in an elderly population. *Brain Behav*. 2020;10:e01662. doi:10.1002/brb3.1662
13. van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in Early Alzheimer's Disease. *N Engl J Med*. 2023;388:9-21. doi:10.1056/NEJMoa2212948

14. Sevigny J, Chiao P, Bussière T, et al. The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. *Nature*. 2016;537:50-56. doi:10.1038/nature19323
15. Tolar M, Abushakra S, Hey JA, Porsteinsson A, Sabbagh M. Aducanumab, gantenerumab, BAN2401, and ALZ-801-the first wave of amyloid-targeting drugs for Alzheimer's disease with potential for near term approval. *Alzheimers Res Ther*. 2020;12:95. doi:10.1186/s13195-020-00663-w
16. Schneider L. A resurrection of aducanumab for Alzheimer's disease. *Lancet Neurol*. 2020;19:111-112. doi:10.1016/S1474-4422(19)30480-6
17. Friedlinger M, Schad LR, Blüml S, Tritsch B, Lorenz WJ. Rapid automatic brain volumetry on the basis of multispectral 3D MR imaging data on personal computers. *Comput Med Imaging Graph*. 1995;19:185-205. doi:10.1016/0895-6111(94)00045-e
18. Kitagaki H, Mori E, Yamaji S, et al. Frontotemporal dementia and Alzheimer disease: evaluation of cortical atrophy with automated hemispheric surface display generated with MR images. *Radiology*. 1998;208:431-439. doi:10.1148/radiology.208.2.9680572
19. Fischl B. FreeSurfer. *Neuroimage*. 2012;62:774-781. doi:10.1016/j.neuroimage.2012.01.021
20. Sargolzaei S, Sargolzaei A, Cabrerizo M, et al. Estimating intracranial volume in brain research: An evaluation of methods. *Neuroinformatics*. 2015;13:427-441. doi:10.1007/s12021-015-9266-5
21. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34:939-944. doi:10.1212/wnl.34.7.939
22. Jessen F, Spottke A, Boecker H, et al. Design and first baseline data of the dzne multicenter observational study on predementia alzheimer's disease (DELCODE). *Alzheimers Res Ther*. 2018;10:15. doi:10.1186/s13195-017-0314-2
23. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*. 2011;134:2456-2477. doi:10.1093/brain/awr179
24. Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. *Neurology*. 2011;76:1006-1014. doi:10.1212/WNL.0b013e31821103e6
25. Rudolph J, Huemmer C, Preuhs A, et al. Non-radiology healthcare professionals significantly benefit from ai-assistance in emergency-related chest radiography interpretation. *Chest*. 2024;166(1):157-170. doi:10.1016/j.chest.2024.01.039
26. Rudolph J, Huemmer C, Ghesu F-C, et al. Artificial intelligence in chest radiography reporting accuracy: added clinical value in the emergency unit setting without 24/7 radiology coverage. *Invest Radiol*. 2022;57:90-98. doi:10.1097/RLI.0000000000000813
27. Rudolph J, Fink N, Dinkel J, et al. Interpretation of thoracic radiography shows large discrepancies depending on the qualification of the physician-quantitative evaluation of interobserver agreement in a representative emergency department scenario. *Diagnostics (Basel)*. 2021;11:1868. doi:10.3390/diagnostics11101868
28. Rueckel J, Trappmann L, Schachtner B, et al. Impact of confounding thoracic tubes and pleural dehiscence extent on artificial intelligence pneumothorax detection in chest radiographs. *Invest Radiol*. 2020;55:792-798. doi:10.1097/RLI.0000000000000707
29. Rueckel J, Kunz WG, Hoppe BF, et al. Artificial intelligence algorithm detecting lung infection in supine chest radiographs of critically ill patients with a diagnostic accuracy similar to board-certified radiologists. *Crit Care Med*. 2020;48:e574-e583. doi:10.1097/CCM.0000000000004397
30. R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria; 2023. https://www.R-project.org/
31. Ronneberger O, Fischer P, Brox T. *U-Net: convolutional networks for biomedical image segmentation*. Springer; 2015.
32. The Royal College of Radiologists. Clinical Radiology Workforce Census 2022 Report. n.d. The Royal College of Radiologists ;2022.
33. Jessen F, Wolfsgruber S, Kleineindam L, et al. Subjective cognitive decline and stage 2 of Alzheimer disease in patients from memory centers. *Alzheimers Dement*. 2022 ;19(2):487-497 doi:10.1002/alz.12674
34. Düzel E, Ziegler G, Berron D, et al. Amyloid pathology but not APOE ε4 status is permissive for tau-related hippocampal dysfunction. *Brain*. 2022;145:1473-1485. doi:10.1093/brain/awab405
35. Wolfsgruber S, Kleineidam L, Weyrauch A-S, et al. Relevance of subjective cognitive decline in older adults with a first-degree family history of alzheimer's disease. *J Alzheimers Dis*. 2022;87:545-555. doi:10.3233/JAD-215416
36. Brosseron F, Maass A, Kleineidam L, et al. Soluble TAM receptors sAXL and sTyro3 predict structural and functional protection in Alzheimer's disease. *Neuron*. 2022;110:1009-1022.e4. doi:10.1016/j.neuron.2021.12.016
37. Teipel SJ, Dyrba M, Ballarini T, et al. Association of cholinergic basal forebrain volume and functional connectivity with markers of inflammatory response in the alzheimer's disease spectrum. *J Alzheimers Dis*. 2022;85:1267-1282. doi:10.3233/JAD-215196
38. Jessen F, Amariglio RE, Buckley RF, et al. The characterisation of subjective cognitive decline. *Lancet Neurol*. 2020;19:271-278. doi:10.1016/S1474-4422(19)30368-0
39. Yu AC, Mohajer B, Eng J. External validation of deep learning algorithms for radiologic diagnosis: a systematic review. *Radiol Artif Intell*. 2022;4:e210064. doi:10.1148/ryai.210064
40. Kuo RYL, Harrison C, Curran T-A, et al. Artificial intelligence in fracture detection: A systematic review and meta-analysis. *Radiology*. 2022;304:50-62. doi:10.1148/radiol.211785
41. Hogan DB, Jetté N, Fiest KM, et al. The prevalence and incidence of frontotemporal dementia: a systematic review. *Can J Neurol Sci*. 2016;43(1):S96-109. doi:10.1017/cjn.2016.25

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Rudolph J, Rueckel J, Döpfert J, et al. Artificial intelligence–based rapid brain volumetry substantially improves differential diagnosis in dementia. *Alzheimer's Dement*. 2024;16:e70037. https://doi.org/10.1002/dad2.70037