


## ORIGINAL ARTICLE

## Regional MRI volumetry using NeuroQuant versus visual rating scales in patients with cognitive impairment and dementia

Karin Persson<sup>1,2</sup>  | Maria L. Barca<sup>1,2</sup> | Trine Holt Edwin<sup>2</sup> | Lena Cavallin-Eklund<sup>3</sup> | Gro Gujord Tangen<sup>1,2,4</sup> | Hanneke F. M. Rhodius-Meester<sup>2,5,6,7</sup> | Geir Selbæk<sup>1,2,8</sup> | Anne-Brita Knapskog<sup>2</sup> | Knut Engedal<sup>1,2</sup>

<sup>1</sup>The Norwegian National Centre for Ageing and Health, Vestfold Hospital Trust, Tønsberg, Norway

<sup>2</sup>Department of Geriatric Medicine, Department of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway

<sup>3</sup>Karolinska Institutet, Stockholm, Sweden

<sup>4</sup>Department of Rehabilitation Science and Health Technology, Faculty of Health Science, Oslo Metropolitan University, Oslo, Norway

<sup>5</sup>Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

<sup>6</sup>Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands

<sup>7</sup>Department of Internal Medicine, Geriatric Medicine Section, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands

<sup>8</sup>Faculty of Medicine, University of Oslo, Oslo, Norway

## Correspondence

Karin Persson, The Norwegian National Centre for Ageing and Health, Vestfold Hospital Trust, Ageing and Health, Postboks 2136, 3103 Tønsberg, Norway. Email: [Karin.Persson@aldringoghelse.no](mailto:Karin.Persson@aldringoghelse.no)

## Funding information

The Southern and Eastern Norway Regional Health Authority, Grant/Award Number: 2018078; Memorabel Dementia Fellowship 2021, Grant/Award Number: 10510022110004; Horizon 2022 project PROMINENT, Grant/Award Number: 101112145; Alzheimer Nederland InterACT grant, Grant/Award Number: WE.08-2022-06

## Abstract

**Background and purpose:** The aims were to compare the novel regional brain volumetric measures derived by the automatic software NeuroQuant (NQ) with clinically used visual rating scales of medial temporal lobe atrophy (MTA), global cortical atrophy-frontal (GCA-f), and posterior atrophy (PA) brain regions, assessing their diagnostic validity, and to explore if combining automatic and visual methods would increase diagnostic prediction accuracy.

**Methods:** Brain magnetic resonance imaging (MRI) examinations from 86 patients with subjective and mild cognitive impairment (i.e., non-dementia,  $n = 41$ ) and dementia ( $n = 45$ ) from the Memory Clinic at Oslo University Hospital were assessed using NQ volumetry and with visual rating scales. Correlations, receiver operating characteristic analyses calculating area under the curves (AUCs) for diagnostic accuracy, and logistic regression analyses were performed.

**Results:** The correlations between NQ volumetrics and visual ratings of corresponding regions were generally high between NQ hippocampi/temporal volumes and MTA ( $r = -0.72/-0.65$ ) and between NQ frontal volume and GCA-f ( $r = -0.62$ ) but lower between NQ parietal/occipital volumes and PA ( $r = -0.49/-0.37$ ).

AUCs of each region, separating non-dementia from dementia, were generally comparable between the two methods, except that NQ hippocampi volume did substantially better than visual MTA (AUC = 0.80 vs. 0.69). Combining both MRI methods increased

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Brain and Behavior* published by Wiley Periodicals LLC.

only the explained variance of the diagnostic prediction substantially regarding the posterior brain region.

**Conclusions:** The findings of this study encourage the use of regional automatic volumetry in locations lacking neuroradiologists with experience in the rating of atrophy typical of neurodegenerative diseases, and in primary care settings.

#### KEYWORDS

automatic regional volumetry, cognitive impairment, dementia, MRI, NeuroQuant, visual rating scale

## 1 | INTRODUCTION

The need for efficient and cost-effective diagnostic tools to assist in the clinical evaluation of cognitive decline and its underlying etiologies is urgent due to the rising prevalence of dementia and the potential availability of new treatments for Alzheimer's disease (AD) in the near future (Perneczky et al., 2023; World Health Organization, 2022). Structural magnetic resonance imaging (MRI) of the brain has gained increased relevance during the last years and is now not only used to exclude non-dementia etiologies but also to support in finding evidence for disease-specific atrophy patterns. While visual rating scales to evaluate regional structural changes of the most relevant brain regions have been in use for the last decades (Koedam et al., 2011; Pasquier et al., 1996; Scheltens et al., 1992), the scales are subjective and rather coarse. Visual ratings of the frontal (global cortical atrophy-frontal scale, GCA-f) and posterior (posterior atrophy scale, PA) brain regions are known to be challenging with lower intra- and interrater reliability compared to ratings of medial temporal lobe atrophy (MTA), and with less diagnostic value of the PA for the oldest patients (Ferreira et al., 2015; Koedam et al., 2011; Rhodius-Meester et al., 2017). In addition, as many as one-third of radiologists have been found to lack complete confidence in the visual rating methods (Vernooij et al., 2019). Support of automatic quantification methods might aid in this challenge. Thus, automatic and data-driven methods have been developed and are widely used in research but to a lesser extent incorporated into everyday clinical routine (Pemberton et al., 2021). The adoption of commercially available clinically feasible automatic methods in clinical practice might be hampered by their insufficient clinical validation. Indeed, a 2021 review found that some methods had been validated against manual segmentation and visual rating scales, but none of the included methods presented evidence as to how the tools should be integrated into clinical settings or how they should contribute to the diagnostic workup (Pemberton et al., 2021).

At the Memory Clinic at Oslo University Hospital (OUH), Norway, the automatic volumetric method NeuroQuant (NQ) (Brewer et al., 2009) has been in use since 2009. During the first years of use, only a limited number of brain regions were assessed by the method. In previous studies, the hippocampus volume as reported by NQ and visual ratings with the MTA scale have been found to correlate well and have equivalent discriminatory power in separating AD from non-

dementia patients (Min et al., 2017; Persson et al., 2018). Currently, newer versions of the NQ software produce volumetric measures of more than 50 brain regions, including the clinically relevant frontal, parietal, occipital, and temporal regions (NeuroQuant, n.d.). To the best of our knowledge, no previous NQ study has validated these novel regional measures.

The main aim of the present study was to compare the novel regional volumetric measures of NQ with the well-known and clinically validated visual rating scales through correlation analyses and to compare their diagnostic accuracy and explore if combining both methods would add value to diagnostic prediction.

## 2 | METHODS

### 2.1 | Participants

Participants were recruited from the Memory Clinic at OUH and had been enrolled in The Norwegian Registry of Persons Assessed for Cognitive Symptoms (NorCog) with written consent for their data in the registry to be used for research (Medbøen et al., 2022). All patients received standardized workup and were referred to brain MRI according to clinical indication. For the present study, we included patients who had been visually rated by an experienced neuroradiologist (Lena Cavallin-Eklund) as part of a previous study (Persson et al., 2017). Of these 218 patients, 88 had been scanned at the research MRI scanner at OUH, where the NQ software was installed. Selection for referral to this specific research scanner, and not to another MRI scanner, was done at random, mainly based on availability and capacity. Of the 88 patients, two were categorized as healthy and were excluded from the sample.

### 2.2 | Diagnoses

The 88 patients were diagnosed according to clinical diagnostic criteria by two physicians (Karin Persson and Trine Holt Edwin) retrospectively based on all clinical information from NorCog and the medical records from the clinical examination (Medbøen et al., 2022). The Jessen criteria were used to diagnose subjective cognitive decline (SCD,  $n = 21$ ),

and NIA/AA criteria were used to diagnose mild cognitive impairment (MCI,  $n = 20$ ) and dementia ( $n = 45$ ) (Albert et al., 2011; Jessen et al., 2014). The majority of the patients with dementia had AD according to the NIA/AA criteria (either probable AD or possible AD mixed with other pathology,  $n = 34$ ) (Albert et al., 2011). To diagnose other dementia etiologies, the following criteria were used: Rascovsky et al. (2011) for frontotemporal dementia (FTD,  $n = 2$ ), McKeith et al. (2017) for dementia with Lewy bodies (DLB,  $n = 3$ ), Emre et al. (2007) for Parkinson disease dementia (PDD,  $n = 3$ ), and VASCOG criteria for vascular dementia (VaD,  $n = 1$ ) (Sachdev et al., 2014). Additionally, two patients had unspecified dementia. Patients with SCD or MCI are hereafter referred to as “non-dementia”.

### 2.3 | MRI acquisition and analysis

All MRIs were performed on the same GE SIGNA HDxt 3T scanner (GE Healthcare). The scans were rated by an experienced neuroradiologist (Lena Cavallin-Eklund) and analyzed with NQ (version 3, CorTechs Labs/University of California) (Brewer et al., 2009).

Visual ratings of the medial temporal, frontal, and posterior brain lobes were carried out according to three well-validated scales using T1-sequences. The Scheltens scale was used to rate the atrophy of the medial temporal lobe (MTA) on a scale from zero to four (Scheltens et al., 1992). The mean score of the left and right side was used in this study. The GCA-f scale was used to rate the global cortical atrophy of the frontal lobes on a scale from zero to three (Pasquier et al., 1996), and the PA scale was used to rate cortical atrophy of the posterior brain region (mainly parietal lobes and parietooccipital sulcus) on a scale from zero to three (Koedam et al., 2011). On all three scales, a higher score indicates more atrophy.

NQ produces volumetric measures of several brain regions including the hippocampus as well as joint volumetric measures of the temporal, frontal, parietal, and occipital cortical regions. Volumes are presented as both raw volumes and proportions of intracranial volume (ICV), as well as percentiles calculated based on data from healthy individuals (CorTechs Labs., Inc, n.d.). In this study, we report regional volumes as the total volume of both hemispheres as a proportion of ICV. We used hippocampi and temporal volumes as the NQ correlates to MTA, frontal volume as the NQ correlates to GCA-f, and parietal and occipital volumes as the NQ correlates to PA.

### 2.4 | Statistics

Analyses were carried out using IBM SPSS Statistics for Windows (version 27, Armonk) with a significance level set at .05. Independent samples  $t$ -test and  $\chi^2$  tests were used for descriptive comparisons. Pearson correlation coefficients were produced to compare the visual and volumetric normal distributed measures, interpreted according to Cohen, that is,  $r < 0.1$ —very small,  $0.1 \leq r < 0.3$ —small,  $0.3 \leq r < 0.5$ —moderate, and  $r \geq 0.5$ —large (Cohen, 1988). Receiver operating characteristics (ROC) analyses, calculating the area under the

curve (AUC), were performed to validate the performance of each MRI classifier (i.e., the three visual measures and the five NQ volumetrics) in distinguishing dementia from non-dementia (both SCD and MCI) and from SCD alone. An AUC is generally regarded as poor if  $0.5 < \text{AUC} < 0.7$ , acceptable if  $0.7 \leq \text{AUC} < 0.8$ , excellent if  $0.8 \leq \text{AUC} < 0.9$ , and outstanding if  $\geq 0.9$  (Hosmer et al., 2013). To investigate the potential synergistic value of combining both visual and NQ volumetric methods in enhancing diagnostic prediction, we conducted three logistic regression analyses. These three analyses used dementia/non-dementia as the outcome variable, with each analysis focusing on one specific brain region. Model a included visual measures, model b included automatic measures, and model c included both visual and automatic measures. The risk of multicollinearity increases when correlated variables with  $r$  above 0.5 or 0.8 are included in the same model (Shrestha, 2020), and the correlation coefficients between MRI measures of each region varied between  $-0.37$  and  $-0.72$  (Table 2). However, as the main purpose of including correlated variables in the same model was to compare the explained variance of each model using Nagelkerke  $R^2$ , the risk of multicollinearity, especially relevant for the analysis of the temporal measures ( $r = -0.72$ ), was taken into consideration but regarded acceptable. All regression models were adjusted for age and sex.

### 2.5 | Ethics

All patients gave written informed consent to be included in NorCog including that all information collected at their examinations at the clinic, including supplemental data such as MRI results, can be used for research. The present study was approved by the Regional Committee of Medical Research Ethics of the South-East Norway Regional Health Authority (REC South-East number 2019/79).

## 3 | RESULTS

Patient characteristics are presented in Table 1. The mean (SD) age was 71.9 (8.0) in the dementia group and 67.9 (10.1) in the non-dementia group ( $p = .046$ ), and 22 (49%) versus 14 (34%) were females ( $p = .166$ ). Patients with dementia had significantly more atrophy on all MRI measures compared to the non-dementia patients ( $p \leq .003$ ).

Correlations between all MRI measures are presented in Table 2. The highest correlation was found between the NQ volumetry of the hippocampi and the visual rating of MTA ( $r = -0.72$ ). For all regional NQ volumetrics, the correlation was highest with the visual rating scale of the corresponding region, that is, between NQ hippocampi and temporal volumes and the MTA ( $r = -0.72/-0.65$ ), between NQ frontal volume and the GCA-f ( $r = -0.62$ ), and between NQ parietal and occipital volumes and the PA ( $r = -0.49/-0.37$ ).

To validate the performance of the different MRI classifiers, ROC analyses were performed using dementia versus non-dementia as the outcome (Table 3, column a). The AUCs of the NQ hippocampi and temporal volumes were 0.74 and 0.80 (“acceptable/excellent”),

**TABLE 1** Patient characteristics.

	SCD (n = 21)	MCI (n = 20)	Dementia (n = 45)
Age (years)	65.2 (9.7)	70.8 (9.9)	71.9 (8.0)
Females, (n, %)	7 (33%)	7 (35%)	22 (49%)
Education (years)	14.5 (3.5)	14.2 (3.2)	12.9 (3.6)
MMSE (score)	29.3 (0.7)	26.7 (2.9)	22.3 (5.8)
CDR-SB (score)	0.3 (0.4)	1.7 (1.0)	5.8 (3.2)
MTA mean (score)	1.1 (0.6)	1.5 (0.9)	2.0 (1.0)
GCA-f (score)	0.3 (0.5)	0.6 (0.8)	0.8 (0.6)
PA (score)	0.3 (0.6)	0.4 (0.1)	1.0 (0.8)
Hippocampi/ICV (%)	0.51 (0.07)	0.42 (0.09)	0.40 (0.08)
Temporal/ICV (%)	8.5 (0.7)	8.1 (0.9)	7.3 (1.0)
Frontal/ICV (%)	11.5 (0.9)	11.0 (1.4)	10.5 (0.9)
Parietal/ICV (%)	7.1 (0.6)	6.9 (0.8)	6.4 (0.8)
Occipital/ICV (%)	3.9 (0.3)	3.8 (0.4)	3.5 (0.4)

Note: All continuous variables are expressed as mean (SD).

Abbreviations: CDR-SB, clinical dementia rating scale sum of boxes; GCA-f, global cortical atrophy-frontal; ICV, intracranial volume; MCI, mild cognitive impairment; MMSE, mini mental status examination; MTA, medial temporal lobe atrophy; PA, posterior atrophy; SCD, subjective cognitive decline.

**TABLE 2** Correlations between NQ volumetric measures and visual rating measures.

NQ volumetric measures	Visual rating measures		
	MTA mean	GCA-f	PA
Hippocampi/ICV	-0.72	-0.49	-0.34
Temporal/ICV	-0.65	-0.47	-0.39
Frontal/ICV	-0.60	-0.62	-0.44
Parietal/ICV	-0.38	-0.41	-0.49
Occipital/ICV	-0.32	-0.21*	-0.37

Note: Bold values indicate a correlation between measures from equivalent regions of the two measures. Pearson  $r$ ,  $p < .01$  for all except \* $p .059$ .

Abbreviations: GCA-f, global cortical atrophy-frontal; ICV, intracranial volume; MTA, medial temporal lobe atrophy; PA, posterior atrophy.

respectively, and the AUC of MTA was 0.69 ("poor"); the AUC of NQ frontal volume was 0.70 ("acceptable"), and the AUC of GCA-f was 0.69 ("poor"); finally, the AUCs of NQ parietal and occipital volumes were 0.74 and 0.73, respectively, and the AUC of PA was 0.72 ("acceptable"). When using dementia versus SCD as the outcome, all AUCs were within the "acceptable" range, except that NQ hippocampi and temporal volumes increased to "excellent" (0.86 and 0.85) (Table 3, column b). Conjoining all five NQ volumes gave an AUC of 0.77 for dementia versus non-dementia and 0.84 for dementia versus SCD.

In line with the ROC results, the explained variances were higher in the logistic regression model b (NQ volumetric measures) than in model a (visual measures) for all three brain regions (Tables 4–6). The Nagelkerke  $R^2$  of model c (including both NQ volumetric and visual measures) reached a higher level than that of models a and b only regarding the posterior brain region.

## 4 | DISCUSSION

The correlations between regional automatic NQ volumetric measures and corresponding visual rating scale measures were generally high, with the highest correlation found between the temporal NQ volumes and MTA, and the lowest between the posterior NQ volumes and PA. The discriminatory power of each regional NQ volumetric measure and its corresponding visual rating measure were comparable, except that the NQ volumetry of the hippocampi and temporal regions was substantially better at discriminating dementia from non-dementia compared to MTA. Combining NQ volumetrics with visual rating measures increased the diagnostic prediction accuracy in posterior brain regions.

In line with previous studies, the highest correlation was found between measures of the medial temporal region. Previous studies have found automatic quantification methods of the medial temporal region to correlate well with both manual and visual measurements (Koikkalainen et al., 2019; Mårtensson et al., 2020), while correlations for frontal regions have been found to be substantially lower (Koikkalainen et al., 2019; Mårtensson et al., 2019). In a study by Koikkalainen et al. (2019), the medial temporal region correlation coefficients were 0.83 and 0.78, substantially higher than ours, and the coefficient for the global cortical atrophy was 0.64, similar to our NQ frontal volume/GCA-f results. They did not study posterior regions exclusively. To our knowledge, no study using a clinically feasible automatic software has presented data on the parietal or posterior brain regions. The automatic visual ratings of the atrophy model, which is not yet clinically available, present results for the parietal region that are in line with those of the medial temporal region (Mårtensson et al., 2019), that is, higher than the ones we found for the posterior region.

While few previous results using clinically feasible software are available, our current results showing that NQ volumetrics correlate

**TABLE 3** Discriminatory properties. (a) Dementia versus non-dementia and (b) dementia versus subjective cognitive decline (SCD).

	(a)		(b)	
	AUC (95% CI)	<i>p</i>	AUC (95% CI)	<i>p</i>
Visual measures				
MTA mean	0.69 (0.58–0.81)	.002	0.74 (0.62–0.85)	.002
GCA-f	0.69 (0.58–0.81)	.002	0.74 (0.61–0.87)	.002
PA	0.72 (0.61–0.83)	<.001	0.74 (0.61–0.87)	.002
NQ volumetrics				
Hippocampi/ICV	0.74 (0.63–0.85)	<.001	0.86 (0.77–0.95)	<.001
Temporal/ICV	0.80 (0.71–0.89)	<.001	0.85 (0.76–0.94)	<.001
Frontal/ICV	0.70 (0.60–0.81)	.002	0.77 (0.65–0.89)	<.001
Parietal/ICV	0.74 (0.63–0.84)	<.001	0.78 (0.67–0.90)	<.001
Occipital/ICV	0.73 (0.63–0.84)	<.001	0.79 (0.68–0.90)	<.001

ROC analyses.

Abbreviations: AUC, area under the curve; CI, confidence interval; GCA-f, global cortical atrophy-frontal; ICV, intracranial volume; MTA, medial temporal lobe atrophy; NQ, NeuroQuant; PA, posterior atrophy.

**TABLE 4** Associations with dementia/non-dementia. Models a, b, and c include temporal measures.

	Unadjusted		Adjusted model a		Adjusted model b		Adjusted model c	
	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>
MTA mean	2.4 (1.4–4.1)	.002	2.4 (1.2–4.5)	.010			1.3 (0.6–2.8)	.580
Hippocampi/ICV	0.6 × 10 <sup>-4</sup> (1.9 × 10 <sup>-7</sup> – 1.7 × 10 <sup>-2</sup> )	<.001			0.001 (3.4 × 10 <sup>-8</sup> –17.7)	.162	0.003 (4.8 × 10 <sup>-8</sup> –157.1)	.291
Temporal/ICV	0.3 (0.2–0.5)	<.001			0.3 (0.1–0.6)	.002	0.3 (0.1–0.6)	.003
Nagelkerke R <sup>2</sup>				.21		.40		.40

Note: Logistic regression analyses. Models a, b, and c are adjusted for age and sex.

Abbreviations: CI, confidence interval; ICV, intracranial volume; MTA, medial temporal lobe atrophy; OR, odds ratio.

**TABLE 5** Associations with dementia/non-dementia. Models a, b, and c include frontal measures.

	Unadjusted		Adjusted model a		Adjusted model b		Adjusted model c	
	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>
GCA-f	3.0 (1.4–6.5)	.005	2.7 (1.1–6.3)	.025			1.8 (0.7–4.8)	.214
Frontal/ICV	0.5 (0.3–0.8)	.003			0.5 (0.3–0.8)	.010	0.5 (0.3–1.0)	.060
Nagelkerke R <sup>2</sup>				.17		.20		.22

Note: Logistic regression analyses. Models a, b, and c are adjusted for age and sex.

Abbreviations: CI, confidence interval; GCA-f, global cortical atrophy-frontal; ICV, intracranial volume; OR, odds ratio.

**TABLE 6** Associations with dementia/non-dementia. Models a, b, and c include posterior measures.

	Unadjusted		Adjusted model a		Adjusted model b		Adjusted model c	
	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>
PA	3.2 (1.6–6.3)	<.001	3.2 (1.6–6.4)	.001			2.2 (1.0–4.8)	.038
Parietal/ICV	0.3 (0.2–0.6)	<.001			0.4 (0.1–0.9)	.026	0.5 (0.2–1.2)	.123
Occipital/ICV	0.1 (0.02–0.3)	<.001			0.2 (0.04–1.1)	.063	0.2 (0.04–1.2)	.086
Nagelkerke R <sup>2</sup>				.27		.34		.39

Note: Logistic regression analyses. Models a, b, and c are adjusted for age and sex.

Abbreviations: CI, confidence interval; ICV, intracranial volume; OR, odds ratio; PA, posterior atrophy.



best with their corresponding regional visual rating scale indicate good validity of the NQ volumetric method. Indeed, for all regional NQ volumetrics, the correlation was the highest with the corresponding visual measure. The relatively lower correlation between the posterior measurements could be related to the fact that this region is the most difficult to rate visually (Rhodius-Meester et al., 2017). However, it could also be potentially caused by less variation in the volumetric and visual rating scale scores of posterior regions related to the majority of the included dementia patients suffering from AD, typically involving mainly the medial temporal region, especially in the early phases of the disease. On the contrary, our patients were young, many suffering from young onset AD where posterior cortical atrophy is often present (Koedam et al., 2010).

All AUCs were comparable or better using NQ volumetry versus visual ratings, indicative of a better ability of NQ to discriminate dementia from non-dementia. In a previous study including a subsample of the present cohort, we found NQ volumetry of the hippocampus to be better at discriminating AD dementia from non-dementia (Persson et al., 2018), but no previous study has assessed the validity of the frontal and posterior volumetrics of NQ. In the study by Koikkalainen et al. (2016), another automatic MRI quantification algorithm combining several automatic quantification methods was notably better (higher diagnostic accuracy for AD, FTD, VaD, DLB, and HC) than visual MRI ratings alone.

While NQ performed comparable or better than visual assessments, combining both methods only seemed to affect the prediction accuracy of the posterior brain region. In a study on the effect of adding an automatic volumetry report (Geodesic Information Flows) to a regular radiological evaluation, the diagnostic accuracy (of separating healthy controls, AD and FTD) improved when radiologists were presented with the results of the volume report. However, that study did not analyze the use of the QReport alone (Pemberton et al., 2021). In another study, on separating AD from FTD ( $n = 42$ ), presenting only automatic reports to the radiologist decreased the diagnostic accuracy compared to using visual measures alone, while combining the two increased the accuracy (Vernooij et al., 2018). Overall, it seems the regional NQ volumetrics performed quite comparable to other automatic measures, perhaps somewhat lower in temporal regions, but studies were not comparable in the sample size. While the correlation was lower between posterior measures, and the atrophy of this region being difficult to rate, it was of special interest to find that combining posterior volumetrics to the visual measurement added precision to the prediction of the diagnostic stage.

In the near future, in parallel with, and as a consequence of, the development of the disease-modifying AD treatment, the need of biomarkers specific to the AD pathological processes will presumably change the role of MRI from being a supporting diagnostic tool to being a tool to aid in diagnostic stratification, prognostic prediction, and therapeutic effect evaluation. Simultaneously, the number of patients needing assessments will increase dramatically (World Health Organization, 2022). Thus, the integration of time-efficient automated tools with fewer challenges concerning the need of neuroradiological expertise and lower interrater reliability (Mårtensson et al., 2020) will be

prompted. We believe the present findings support an increased use of automatic software, but further validations in larger and longitudinally followed-up cohorts are necessary.

The main limitation of this pilot study is the sample size. However, to the best of our knowledge, it is the first study comparing the clinically important regional volumetrics of NQ against visual rating measures. Patients were diagnosed retrospectively and without the inclusion of biological biomarkers (i.e., beta-amyloid or phosphorylated tau markers), but this was not regarded as a major drawback for the purpose of validating the method.

To conclude, we found the novel regional NQ volumetrics to correlate well with validated visual rating scales, to have a similar or even better diagnostic discriminatory power, and to add precision to the diagnostic prediction (posterior volumetric measures). We believe these results should be replicated in a larger sample, to ultimately conclude regarding a beneficial use of automatic volumetry in clinical practice.

## AUTHOR CONTRIBUTIONS

**Karin Persson:** Conceptualization; data curation; formal analysis; writing—original draft; writing—review and editing. **Maria Barca:** Data curation; writing—review and editing. **Trine Holt Edwin:** Data curation; writing—original draft; writing—review and editing. **Lena Cavallin-Eklund:** Data curation; investigation; writing—original draft; writing—review and editing. **Gro Gujord Tangen:** Methodology; writing—original draft; writing—review and editing. **Hanneke F.M. Rhodius-Meester:** Methodology; writing—original draft; writing—review and editing. **Geir Selbæk:** Project administration; resources; writing—original draft; writing—review and editing. **Anne-Brita Knapskog:** Investigation; methodology; writing—original draft; writing—review and editing. **Knut Engedal:** Investigation; writing—original draft; writing—review and editing.

## ACKNOWLEDGMENTS

We want to acknowledge the Norwegian Registry of Persons Assessed for Cognitive Symptoms (NorCog) for providing access to patient data. The Southern and Eastern Norway Regional Health Authority supported the study through funding for Karin Persson but was not involved in conducting the study or in the preparation of the manuscript. Hanneke F.M. Rhodius-Meester is supported by the Horizon 2022 project PROMINENT (project number 101112145) and recipient of the Memorabel Dementia Fellowship 2021 (ZonMw project number 10510022110004) and Alzheimer Nederland Inter-ACT grant (project number WE.08-2022-06).

## CONFLICT OF INTEREST STATEMENT

Dr. Persson and Dr. Knapskog report work with the Novo Nordisk NN6535-4730 trial outside the submitted work. Prof. Selbæk reports participation in Roche and Biogen advisory boards. Dr. Rhodius-Meester performs contract research for Combinostics, and all funding is paid to her institution. The other authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## ORCID

Karin Persson  <https://orcid.org/0000-0001-6198-7804>

## PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1002/brb3.3397>.

## REFERENCES

- NeuroQuant. (n.d.). <https://www.cortechs.ai/products/neuroquant/tba/>
- Albert, M. S., Dekosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7, 270–279. <https://doi.org/10.1016/j.jalz.2011.03.008>
- Brewer, J. B., Magda, S., Airriess, C., & Smith, M. E. (2009). Fully-automated quantification of regional brain volumes for improved detection of focal atrophy in Alzheimer disease. *AJNR American Journal of Neuroradiology*, 30, 578–580. <https://doi.org/10.3174/ajnr.A1402>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2 ed.). Lawrence Erlbaum Associates.
- CorTechs Labs. Inc. (n.d.). The NeuroQuant normative database white paper. Statistical power analysis for the behavioral sciences. Our Normative Database - [Cortechs.ai](https://www.cortechs.ai)
- Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., Broe, G. A., Cummings, J., Dickson, D. W., Gauthier, S., Goldman, J., Goetz, C., Kerczyn, A., Lees, A., Levy, R., Litvan, I., Mckeith, I., Olanow, W., Poewe, W., ... Dubois, B. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Movement Disorders*, 22, 1689–1707. <https://doi.org/10.1002/mds.21507>
- Ferreira, D., Cavallin, L., Larsson, E.-M., Muehlboeck, J.-S., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Lovestone, S., Simmons, A., Wahlund, L.-O., & Westman, E. (2015). Practical cut-offs for visual rating scales of medial temporal, frontal and posterior atrophy in Alzheimer's disease and mild cognitive impairment. *Journal of Internal Medicine*, 278, 277–290. <https://doi.org/10.1111/joim.12358>
- Hosmer, Jr. D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
- Jessen, F., Amariglio, R. E., Van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., Dubois, B., Dufouil, C., Ellis, K. A., Van Der Flier, W. M., Glodzik, L., Van Harten, A. C., De Leon, M. J., Mchugh, P., Mielke, M. M., Molinuevo, J. L., Mosconi, L., Osorio, R. S., Perrotin, A., ... Wagner, M. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, 10, 844–852. <https://doi.org/10.1016/j.jalz.2014.01.001>
- Koedam, E. L. G. E., Lauffer, V., Van Der Vlies, A. E., Van Der Flier, W. M., Scheltens, P., & Pijnenburg, Y. A. L. (2010). Early-versus late-onset Alzheimer's disease: More than age alone. *Journal of Alzheimer's Disease*, 19, 1401–1408. <https://doi.org/10.3233/JAD-2010-1337>
- Koedam, E. L. G. E., Lehmann, M., Van Der Flier, W. M., Scheltens, P., Pijnenburg, Y. A. L., Fox, N., Barkhof, F., & Wattjes, M. P. (2011). Visual assessment of posterior atrophy development of a MRI rating scale. *European Radiology*, 21, 2618–2625. <https://doi.org/10.1007/s00330-011-2205-4>
- Koikkalainen, J., Rhodius-Meester, H., Tolonen, A., Barkhof, F., Tijms, B., Lemstra, A. W., Tong, T., Guerrero, R., Schuh, A., Ledig, C., Rueckert, D., Soininen, H., Remes, A. M., Waldemar, G., Hasselbalch, S., Mecocci, P., Van Der Flier, W., & Lötjönen, J. (2016). Differential diagnosis of neurodegenerative diseases using structural MRI data. *NeuroImage: Clinical*, 11, 435–449. <https://doi.org/10.1016/j.nicl.2016.02.019>
- Koikkalainen, J. R., Rhodius-Meester, H. F. M., Frederiksen, K. S., Bruun, M., Hasselbalch, S. G., Baroni, M., Mecocci, P., Vanninen, R., Remes, A., Soininen, H., Van Gils, M., Van Der Flier, W. M., Scheltens, P., Barkhof, F., Erkinjuntti, T., & Lötjönen, J. M. P. (2019). Automatically computed rating scales from MRI for patients with cognitive disorders. *European Radiology*, 29, 4937–4947. <https://doi.org/10.1007/s00330-019-06067-1>
- Mårtensson, G., Ferreira, D., Cavallin, L., Muehlboeck, J.-S., Wahlund, L.-O., Wang, C., & Westman, E. (2019). AVRA: Automatic visual ratings of atrophy from MRI images using recurrent convolutional neural networks. *NeuroImage: Clinical*, 23, 101872. <https://doi.org/10.1016/j.nicl.2019.101872> PMID: 31154242
- Mårtensson, G., Håkansson, C., Pereira, J. B., Palmqvist, S., Hansson, O., Van Westen, D., & Westman, E. (2020). Medial temporal atrophy in preclinical dementia: Visual and automated assessment during six year follow-up. *NeuroImage: Clinical*, 27, 102310. <https://doi.org/10.1016/j.nicl.2020.102310> PMID: 32580125
- Mckeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J.-P., Weintraub, D., Aarsland, D., Galvin, J., Attems, J., Ballard, C. G., Bayston, A., Beach, T. G., Blanc, F., Bohnen, N., Bonanni, L., Bras, J., Brundin, P., Burn, D., Chen-Plotkin, A., ... Kosaka, K. (2017). Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology*, 89, 88–100. <https://doi.org/10.1212/WNL.0000000000004058> PMID: 28592453
- Medbøen, I. T., Persson, K., Nàvik, M., Totland, T. H., Bergh, S., Treviño, C. S., Ulstein, I., Engedal, K., Knapskog, A.-B., Brækhus, A., Øksengård, A. R., Horndalsveen, P. O., Saltvedt, I., Lyngroth, A. L., Ranhoff, A. H., Skrettingland, D. B., Naik, M., Soares, J. Z., Johnsen, B., & Selbaek, G. (2022). Cohort profile: The Norwegian Registry of Persons Assessed for Cognitive Symptoms (NorCog)—A national research and quality registry with a biomaterial collection. *BMJ Open*, 12, e058810. <https://doi.org/10.1136/bmjopen-2021-058810> PMID: 36448543
- Min, J., Moon, W.-J., Jeon, J. Y., Choi, J. W., Moon, Y.-S., & Han, S.-H. (2017). Diagnostic efficacy of structural MRI in patients with mild-to-moderate Alzheimer disease: Automated volumetric assessment versus visual assessment. *American Journal of Roentgenology*, 208, 617–623. <https://doi.org/10.2214/AJR.16.16894>
- Pasquier, F., Leys, D., Weerts, J. G. E., Mounier-Vehier, F., Barkhof, F., & Scheltens, P. (1996). Inter- and intraobserver reproducibility of cerebral atrophy assessment on MRI scans with hemispheric infarcts. *European Neurology*, 36, 268–272. <https://doi.org/10.1159/000117270>
- Pemberton, H. G., Goodkin, O., Prados, F., Das, R. K., Vos, S. B., Moggridge, J., Coath, W., Gordon, E., Barrett, R., Schmitt, A., Whiteley-Jones, H., Burd, C., Wattjes, M. P., Haller, S., Vernooij, M. W., Harper, L., Fox, N. C., Paterson, R. W., Schott, J. M., ... Barkhof, F. (2021). Automated quantitative MRI volumetry reports support diagnostic interpretation in dementia: A multi-rater, clinical accuracy study. *European Radiology*, 31, 5312–5323. <https://doi.org/10.1007/s00330-020-07455-8>
- Pemberton, H. G., Zaki, L. A. M., Goodkin, O., Das, R. K., Steketee, R. M. E., Barkhof, F., & Vernooij, M. W. (2021). Technical and clinical validation of commercial automated volumetric MRI tools for dementia diagnosis—A systematic review. *Neuroradiology*, 63, 1773–1789. <https://doi.org/10.1007/s00234-021-02746-3>
- Pernecky, R., Jessen, F., Grimmer, T., Levin, J., Flöel, A., Peters, O., & Froelich, L. (2023). Anti-amyloid antibody therapies in Alzheimer's disease. *Brain*, 146, 842–849. <https://doi.org/10.1093/brain/awad005>
- Persson, K., Barca, M. L., Cavallin, L., Brækhus, A., Knapskog, A.-B., Selbæk, G., & Engedal, K. (2018). Comparison of automated volumetry of the hippocampus using NeuroQuant(R) and visual assessment of the medial temporal lobe in Alzheimer's disease. *Acta Radiologica*, 59, 997–1001. <https://doi.org/10.1177/0284185117743778> PMID: 29172642

- Persson, K., Barca, M. L., Eldholm, R. S., Cavallin, L., Šaltytė Benth, J., Selbæk, G., Brækhus, A., Saltvedt, I., & Engedal, K. (2017). Visual evaluation of medial temporal lobe atrophy as a clinical marker of conversion from mild cognitive impairment to dementia and for predicting progression in patients with mild cognitive impairment and mild Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *44*, 12–24. <https://doi.org/10.1159/000477342> PMID: 28614836
- Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., Van Swieten, J. C., Seelaar, H., Dopper, E. G. P., Onyike, C. U., Hillis, A. E., Josephs, K. A., Boeve, B. F., Kertesz, A., Seeley, W. W., Rankin, K. P., Johnson, J. K., Gorno-Tempini, M.-L., Rosen, H., ... Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, *134*, 2456–2477. <https://doi.org/10.1093/brain/awr179>
- Rhodijs-Meester, H. F. M., Benedictus, M. R., Wattjes, M. P., Barkhof, F., Scheltens, P., Muller, M., & Van Der Flier, W. M. (2017). MRI visual ratings of brain atrophy and white matter hyperintensities across the spectrum of cognitive decline are differently affected by age and diagnosis. *Frontiers in Aging Neuroscience*, *9*, 117. <https://doi.org/10.3389/fnagi.2017.00117> PMID: 28536518
- Sachdev, P., Kalaria, R., O'Brien, J., Skoog, I., Alladi, S., Black, S. E., Blacker, D., Blazer, D. G., Chen, C., Chui, H., Ganguli, M., Jellinger, K., Jeste, D. V., Pasquier, F., Paulsen, J., Prins, N., Rockwood, K., Roman, G., & Scheltens, P. (2014). Diagnostic criteria for vascular cognitive disorders: A VASCOG statement. *Alzheimer Disease and Associated Disorders*, *28*, 206–218. <https://doi.org/10.1097/WAD.0000000000000034>
- Scheltens, P., Leys, D., Barkhof, F., Huglo, D., Weinstein, H. C., Vermersch, P., Kuiper, M., Steinling, M., Wolters, E. C., & Valk, J. (1992). Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: Diagnostic value and neuropsychological correlates. *Journal of Neurology, Neurosurgery, and Psychiatry*, *55*, 967–972. <https://doi.org/10.1136/jnnp.55.10.967>
- Shrestha, N. (2020). Detecting multicollinearity in regression analysis. *American Journal of Applied Mathematics and Statistics*, *8*, 39–42. <https://doi.org/10.12691/ajams-8-2-1>
- Vernooij, M. W., Jasperse, B., Steketee, R., Koek, M., Vrooman, H., Ikram, M. A., Papma, J., Van Der Lugt, A., Smits, M., & Niessen, W. J. (2018). Automatic normative quantification of brain tissue volume to support the diagnosis of dementia: A clinical evaluation of diagnostic accuracy. *NeuroImage: Clinical*, *20*, 374–379. <https://doi.org/10.1016/j.nicl.2018.08.004>
- Vernooij, M. W., Pizzini, F. B., Schmidt, R., Smits, M., Yousry, T. A., Bargallo, N., Frisoni, G. B., Haller, S., & Barkhof, F. (2019). Dementia imaging in clinical practice: A European-wide survey of 193 centres and conclusions by the ESNR working group. *Neuroradiology*, *61*, 633–642. <https://doi.org/10.1007/s00234-019-02188-y>
- World Health Organization. (2022). *Dementia fact sheet*. <https://www.who.int/news-room/fact-sheets/detail/dementia>

**How to cite this article:** Persson, K., Barca, M. L., Edwin, T. H., Cavallin-Eklund, L., Tangen, G. G., Rhodijs-Meester, H. F. M., Selbæk, G., Knapskog, A.-B., & Engedal, K. (2024). Regional MRI volumetry using NeuroQuant versus visual rating scales in patients with cognitive impairment and dementia. *Brain and Behavior*, *14*, e3397. <https://doi.org/10.1002/brb3.3397>