Journal of Neuroimaging
Official Journal of the American Society of Neuroimaging

Validation of an MRI Rating Scale for Amyloid-Related Imaging Abnormalities

Journal:	*Journal of Neuroimaging*
Manuscript ID	JON-16-4805.R1
Wiley - Manuscript type:	Clinical Investigative Study
Date Submitted by the Author:	30-Nov-2016
Complete List of Authors:	Bechten, Arianne; Image Analysis Centre, Department of Radiology & Nuclear Medicine, VU University Medical Center Wattjes, M; Image Analysis Centre, Department of Radiology & Nuclear Medicine, VU University Medical Center Purcell, Derk; Department of Radiology, California Pacific Medical Center; BioClinica Inc Newark Sanchez Aliaga, Esther; Image Analysis Centre, Department of Radiology & Nuclear Medicine, VU University Medical Center Daams, Marita; Department of Radiology & Nuclear Medicine and Department of Anatomy and Neurosciences, Neuroscience Campus Amsterdam, VU University Medical Center Brashear, H.; Janssen Alzheimer Immunotherapy Research and Development Arrighi, H.; Janssen Alzheimer Immunotherapy Research and Development Barkhof, F; Image Analysis Centre, Department of Radiology & Nuclear Medicine, VU University Medical Center; Institutes of Neurology and healthcare Engineering
Keywords:	Alzheimer's disease (AD), Amyloid Beta (Aβ), Immunotherapy, ARIA (Amyloid-Related Imaging Abnormalities), MRI (Magnetic Resonance Imaging)
Subject Area:	Magnetic Resonance Imaging (MRI) < Magnetic Resonance (MR) < Imaging Techniques < NEUROIMAGING, Brain < Anatomical Region < NEUROIMAGING, Alzheimer-s Disease < Neurodegenerative Diseases < Diseases < NEUROIMAGING

Validation of an MRI Rating Scale for Amyloid-Related Imaging Abnormalities

Arianne Bechten[1], Mike P. Wattjes[1], Derk D. Purcell[2,3] , Esther Sanchez Aliaga[1], Marita Daams[4], H.Robert Brashear[5], H. Michael Arrighi[5], Frederik Barkhof[1,6]

[1] Image Analysis Centre, Department of Radiology & Nuclear Medicine, VU University Medical Center, Amsterdam, the Netherlands

[2] Department of Radiology, California Pacific Medical Center, San Francisco, CA, USA

[3] BioClinica, Newark, CA, USA

[4] Department of Radiology & Nuclear Medicine and Department of Anatomy and Neurosciences, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands

[5] Janssen Alzheimer Immunotherapy Research and Development, Fremont, CA, USA

[6] Institutes of Neurology and healthcare Engineering, UCL, London, UK

Running title: MRI RATING SCALE FOR AMYLOID RELATED IMAGING ABNORMALITIES

Keywords: Alzheimer's disease (AD); Amyloid Beta (Aβ); Immunotherapy; ARIA (Amyloid-Related Imaging Abnormalities); MRI (Magnetic Resonance Imaging).

Please address correspondence to Arianne Bechten, MD, Image Analysis Centre, Department of Radiology & Nuclear Medicine; Room PK -1 x10, VU University Medical Center, De Boelelaan 1118, 1081 HV Amsterdam. The Netherlands; e-mail: a.bechten@vumc.nl.
Phone: +31 20 444 3440

Conflict of interest: We declare that we have not conflict of interest.

Disclosures:

We would like to thank Janssen Alzheimer Immunotherapy Research & Development LLC and Pfizer Inc. for providing the MR imaging scans, collected as part of a phase II study of bapineuzumab IV.

The scientific guarantor of this publication is Prof. Dr. Frederik Barkhof. The authors of this manuscript declare relationship whit the following companies: Dr. Frederik Barkhof acts a consultant for Janssen Alzheimer Immunotherapy Research & Development, LLC and H. Michael Arrighi together with H. Robert Brashear were employees of Janssen Alzheimer Immunotherapy Research & Development, LLC, during the development of this manuscript. All the other authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article. Janssen Alzheimer Immunotherapy Research & Development, LLC, and Pfizer Inc. sponsored this study. One of the authors has significant statistical expertise. Institutional Review Board approval was obtained. Written informed consent was obtained from all subjects (patients) in this study. Methodology: prospective, diagnostic or prognostic study, multicenter study.

Some study subjects or cohorts have been previously reported in “A phase 2 multiple ascending dose trial of bapineuzumab in mild to moderate Alzheimer’s disease”. (Neurology 2009; 73:2061–70); and “Amyloid-related imaging abnormalities in patients with Alzheimer’s disease treated with bapineuzumab: a retrospective analysis”. (Lancet Neurology 2012; 11: 241– 49).

ABSTRACT

Introduction

Immunotherapeutic agents against amyloid β are associated with adverse events, including amyloid-related imaging abnormalities with edema and effusion (ARIA-E). Recently an MRI rating scale was developed for ARIA-E detection and classification. The aim of this study was to validate the use of this rating scale in a larger patient group with multiple raters.

Methods

MRI scans of 75 patients (29 with known ARIA-E and 46 control subjects) were analyzed by 5 neuroradiologists with different degrees of expertise, according to the ARIA-E rating scale. For each patient we included a baseline and a follow-up fluid attenuated inversion recovery (FLAIR) image. Inter-rater agreement was calculated using intraclass correlation coefficient (ICC).

Results

On average 4.1% of the ARIA-E cases were missed. We observed a high inter-rater agreement for scores of sulcal hyperintensity (ICC=0.915; 95% CI, 85-95) and for the combined scores of the two ARIA-E findings, parenchymal and sulcal hyperintensity (ICC=0.878; 95% CI, 79–93). A slightly lower agreement for parenchymal hyperintensity (ICC=0.678; 95% CI, 51-81) was noted.

Conclusion

The ARIA-E rating scale is a simple tool to evaluate the extent of ARIA-E in patients recruited into Aβ-lowering therapeutic trials. It shows high inter-rater agreement among raters with different degrees of expertise.

Introduction

Alzheimer's disease (AD) is a neurodegenerative disease with an increasing prevalence worldwide having a significant socio-economic impact.[1, 2] MRI is one of the most relevant diagnostic tools in the diagnosis of dementia and has been included into current diagnostic criteria.[3] Imaging findings include the demonstration of certain atrophy patterns and the detection of vascular (co) morbidity supporting the clinical diagnosis of AD or other diseases associated with dementia.[4]

Advances in understanding the molecular background of AD pathology strongly support the amyloid cascade hypothesis claiming amyloid β (Aβ) as the main cause for neuronal death and dysfunction.[5, 6] This dysregulated Aβ metabolism and its pathophysiological importance have led to the development of numerous therapeutic approaches to find potential disease-modifying agents targeting Aβ such as active and passive immunotherapy.[5, 7, 8] Due to the nature of AD, patients are often incapable of signaling adverse effects of these new therapeutic strategies. Consequently, the adverse effects of these new therapies often escape clinical detection. Precise monitoring by other means than clinical observation like MRI have become pivotal tools for detecting adverse events in AD trials. Therefore, neuroradiological monitoring plays an important role in therapeutic trials in AD.

MRI is able to identify so-called amyloid-related imaging abnormalities (ARIA) such as cerebral microbleeds and superficial hemosiderosis deposits, designated as ARIA-Hemosiderin (ARIA-H). In addition to these findings related to hemosiderin, MRI

hyperintensities indicative of edema and effusion have been described and named ARIA-Edema and Effusion (ARIA-E).[9] In ARIA-E, three different imaging features can be discerned: parenchymal hyperintensity (PH), sulcal hyperintensity (SH) and gyral swelling (SW).[9, 10] In a minority of cases, these abnormalities were symptomatic primarily associated with headache, confusion, vomiting, or gait disturbance.[9, 11] ARIA-E has been identified most frequently during treatment with amyloid-lowering drugs, especially immunotherapy, but can also occur spontaneously, particularly in patients presenting with cerebral amyloid angiopathy.[9, 12] The clinical consequences of these radiological findings are still unclear; ARIA-E currently remains largely a radiological construct.

For clinical and research purposes, mentioned above, it is imperative to have a robust and validated radiological scoring system that enables the correct radiological classification and grading of ARIA-E, at the time of diagnosis as well as during the treatment. Such a radiological scoring system could assist in the titration of drugs and duration of treatment. As these imaging abnormalities are clinically undetectable in most cases[9, 13] an MRI scale could be a valuable monitoring tool for the detection of ARIA-E events.

Recently, we proposed a rating scale for ARIA-E, which consists of a simple five-point severity rating for each region affected. Six brain areas used for this rating are the same used on the Age-Related White Matter Changes rating scale (ARWMC)[14] because all of these regions can be also affected by ARIA-E. Our initial assessment showed promising inter-rater agreement between two experienced raters in a small sample size (n=10) of ARIA-E patients.[15] So far this scale has not been tested in a group of raters with a more varying neuro-radiological experience and who have not been involved in the scale development. Further testing using multiple raters in a larger dataset was necessary to validate the results from our

initial pilot study; this is mandatory for any rating scale before it can be used with confidence in clinical practice.

The aim of this new study was to validate the use of this novel ARIA-E MRI rating scale in a larger patient group with multiple raters to establish the clinical applicability of the ARIA rating scale.

Methods

Patient Population

All patients included in this study participated in a phase II multicenter, randomized, double-blinded, placebo-controlled, multiple ascending-dose study of bapineuzumab, a humanized anti-amyloid-beta (Aβ) antibody. The study was conducted to evaluate the safety and efficacy of bapineuzumab in 234 patients with mild to moderate AD, with an age range from 50 to 85 years. Patients were randomly assigned to receive placebo or ascending doses (0.15 mg/kg, 0.5mg/kg, 1.0 mg/kg and 2.0 mg/kg) of intravenous bapineuzumab every 13 weeks with up to 6 infusions during 18 months.[13, 16]

Additional inclusion criteria were a mini-mental state examination (MMSE) score of 16-26, where a 16-21 MMSE score was categorized as low and a 22-26 score was high.[11]

Safety scans were performed approximately 6 weeks after each infusion starting at baseline and continuing up to week 71. Imaging included axial FLAIR sequences to detect ARIA-E. MRI was performed on 1.5 T scanners with 5 mm sections obtained in a 2D mode with 1 mm in-plane resolution, as described previously.[16] After study completion, each MRI was re-read centrally by 2 radiologists who identified the presence of ARIA-E in consensus.[13]

For the evaluation sample, we selected 75 patients from the abovementioned patient group with mild to moderate AD. This included all 29 patients with ARIA-E during the bapineuzumab phase-2 program. Patients were not selected on severity or otherwise to avoid the possible impact of selection bias. The majority of ARIA-E cases were first detected during the study and a minority of ARIA-E was first detected after rereading the MRI's

(retrospective analysis). For comparison, we also selected 46 cases without ARIA-E, matched for age, gender and disease severity.

Table 1 shows a detailed summary of the demographic and baseline characteristics of the 75 selected patients. The mean age was 67 years with slightly more females in both groups (overall 62%). Both ARIA-E and non-ARIA-E groups had a mean MMSE score of 21 at enrollment. The ApoE e4 allelic frequencies in the 2 groups were 78% in ARIA-E and 77% in non ARIA-E respectively.

Rater selection and qualification

Five raters were invited to participate in this study. All raters were neuroradiologists; two of the raters had experience in the assessment of ARIA-E and were trained using the rating scale (F.B., M.P.W.).[15] The other three raters had no prior experience using this rating scale and they had different backgrounds in terms of qualifications and years of neuroradiological experience. The average experience in neuroradiology was 10 years, ranging from 4 to 25 years. After instruction and training using an online training module, scoring was performed in a blinded fashion with respect to the clinical information and to the scoring of the other raters.

ARIA-E Scoring

Axial FLAIR scan pairs of each patient, including a baseline and follow-up scan, were presented to each rater digitally. At the pre-study drug baseline MRI scans, no ARIA-E was present. On follow-up MRI scan ARIA-E was present in 29 of the patients. The MRI rating scale for ARIA-E was applied to the FLAIR images using an on-line viewing tool and the results were documented using a digital scoring form.

ARIA-E Scale

The ARIA-E rating process has been described ~~previously [15]. Briefly, for each case of ARIA-E (identified by presence of PH, SH or both) twelve~~previously.[15] The rating scale for ARIA-E included both the location and magnitude of presentation of parenchymal hyperintensities, sulcal hyperintensities and gyral swelling. ARIA-E was defined in accordance with the guidelines of the Alzheimer's association research Round Table Workgroup, including the occurrence of either Sulcal Hyperintensity or Parenchymal Hyperintensity.[9,15] Twelve brain regions are rated separately for each abnormality (PH, SH and SW) (Table 2). A total score is calculated by summing the individual scores from the 6 bilateral regions: frontal lobe, parietal lobe, temporal lobe, occipital lobe, central region (basal ganglia, thalamus, internal and external capsules, corpus callosum, insula), and infratentorial region (brainstem and cerebellum). Each region is scored from 0 to 5 based on the spatial extension and multi-focally of the abnormality. Each item is rated on the presence of absence of ARIA-E abnormality (score 0=normal, score 1=monofocal ≤2 cm, score 2=multifocal ≤2cm, score 3=any lesion >2 but <4cm, score 4=any lesion ≥ 4 cm, and score 5=entire lobe). For each scan, a maximum score (Range, 0 – 60) is derived by summing up the twelve regional scores, using the highest score in each region for each of the three types of abnormalities (PH, SH, SW).[15] The mean and range for the total sum ARIA-E scores were determined for all scores given by the five raters to provide insight into the variation of ARIA-E MRI presentation in our population.

Scale Tutorial and Image Evaluation

Specific instructions on how to perform and grade the ARIA-E changes according to the rating scale were provided using a web-based instructional system. For introduction and

reading purposes the tutorial was designed specifically with 3 separate sections. The first section included a brief introduction into ARIA-E providing the theoretical background and image examples. The second section provided an interactive tutorial on the usage of the rating scale form and 4 training cases, each one with its corresponding baseline and follow-up images, to give the raters a reference for each type of ARIA-E abnormality and how to measure them. The last section was dedicated to the scoring of the data set, which was restricted by a password provided to each rater, and allowed interactive scoring including a measurement tool. Subsequently, the five neuroradiologists had full access to the 75 pair of scans with its corresponding electronic scoring form. All raters were fully blinded to any information regarding treatment and clinical presentation.

Statistical Analysis

To determine agreement between raters the intraclass correlation coefficient (ICC) was calculated for the combination of parenchymal hyperintensity and sulcal hyperintensity together (ARIA-E), followed by parenchymal hyperintensity, sulcal hyperintensity, and swelling separately and the combination of the 3 components. Statistical analyses were conducted using the Statistical Package for the Social Sciences, Version 17, for Windows (SPSS, Chicago, Illinois).

Results

High agreement was observed between the raters regarding the presence or absence of ARIA-E within individual patients. The ICC for the identification of ARIA-E in the group of patients with and ~~controls~~without ARIA-E was 0.953; 95% (93-96). On average, 4.1% of the 29 cases of ARIA-E were missed by each of the 5 raters, an average of 1.2 patients per rater. Three cases of ARIA-E were not identified by at least one rater.

Individual patients presented a wide range of ARIA-E manifestations, with varying severity and different spatial extension. Figures 1, 2 and 3 show the average scores with ranges for minimum and maximum score for each patient for ARIA-E (sum of SH, PH and SW); SH, and PH respectively. Visual inspection showed that the ranges per patient were relatively small for ARIA-E scores and SH scores. Larger ranges in scoring were seen in PH scores. Figures 4 to 7 are examples of cases with discrepant scores.

The range of the ARIA-E score among the 29 scans from the ARIA-E cases was from 1.0 to 44.4, with a mean of 7.4. The range of parenchymal hyperintensity was 0-18.8 (mean 1.2), the range for sulcal hyperintensity was 0-45.4 (mean 4.0), and for swelling the range was 0-34.8 (mean 7.2). Table 3 presents the median scores by rater. The ARIA-E sensitivity and specificity are shown in Table 4, with an average sensitivity of 95.6% and specificity of 98.4%.

Inter-rater agreement (ICC) for sulcal hyperintensity was 0.92 (95% CI 0.85-0.95) and for the combined scores of two ARIA-E findings (parenchymal and sulcal hyperintensity) was 0.88

(95% CI 0.79-0.93). The overall concordance between the 5 raters for PH was (ICC) 0.68; 95% 0.51-0.81) (Table 5). The ICC for SW was 0.66 (95% CI 0.48-0.80).

Discussion

This interobserver study confirms a high agreement in the identification of ARIA-E cases with less than 5% missed cases of ARIA-E per rater. In addition, it verifies a good inter-rater agreement for scores of sulcal hyperintensity and for the combined scores of the two findings used to identify ARIA-E, parenchymal and sulcal hyperintensity, using the ARIA-E scale. Combining the PH score with SH score increased the ICC compared to the ICC of PH alone. These results confirm a preliminary, smaller study in 20 patients, 10 with and 10 without ARIA-E[16] and extend our findings to a larger group of less experienced raters. The range of mean scores confirms that the cases rated represent a wide spectrum of radiological severity. Sensitivity and specificity for detection of ARIA-E cases were high across all five raters. We consider the ARIA-E rating scale a robust and ~~easily applicable tool~~reproducible scale to classify ARIA-E in patients undergoing scheduled MR imaging as a part of drug surveillance.[9, 10, 13]

The ICC was numerically higher than in the previous study, perhaps indicating that SH may be more recognizable or that SH is the most common MRI imaging feature of ARIA-E.[10] The relatively lower ICC for PH across the raters in our study is in agreement with the previous reports in a smaller study.[15] This may reflect over or under-scoring of the largest cross-sectional measurement of the PH, or that the multifocality of PH abnormality was not completely taken into account in the total scoring.

The pathophysiological concept of ARIA is based on an increased vascular permeability associated with Aβ removal from cerebral blood vessels presumably related to amyloid clearance.[9, 13, 17] In both ARIA-H and ARIA-E, the pathophysiology may result from leakage

of intravascular contents due to a shift in amyloid. In particular with ARIA-E, the leakage or effusion of proteinaceous fluid is believed to induce MRI changes suggestive of edema and/or sulcal effusions seen as high signal intensities on ~~fluid-attenuated inversion recovery images (~~FLAIR images.[9, 11, 16] Although swelling often accompanies SH, its detection remains more of a challenge due to the often subtle nature of the swelling as shown by the relatively low ICC compared to the other manifestations. In the current study, the ICC for swelling was slightly higher when compared to the previously reported value. Interestingly, the ICC of the combined ARIA-E score (PH+SH) is closer to the SH than the PH value. Probably, raters were more confident about detecting any hyperintensity, than about the individual contributing elements.

Differentiation between sulcal and parenchymal hyperintensities as separate manifestations remains a challenge. The determination of the extension and boundaries for PH may be difficult.[10] Consequently, a possible limitation of the ARIA-E rating scale, despite its relatively simple nature, is that it requires good neuroanatomical knowledge and its consistent application. We consider the scale to be best suited for experienced neuroradiologists. In addition, we believe that to ensure high inter-observer reliability in rating ARIA-E cases the sum score should be used.

Formatted: Underline color: Auto

Several potential limitations to our study should be considered. Challenges in the assessment of ARIA-E stem from difficulties to detect and score ARIA-E when there are pre-existing vascular white matter lesions within areas of PH, or when (predominantly infratentorial) flow artifacts[12] or impaired image quality produce false-positive findings. Figure 6 shows an example of the difficulties to differentiate between vascular white matter lesions and PH. Nonetheless, as mentioned above, sensitivity and specificity remained high across all five

readers. A second possible limitation of our study is the fact that our dataset is still relatively small and relatively young. However this is the largest dataset of ARIA-E patients in which inter-rater reliability has been studied prospectively. Another possible limitation is that the frequency or ARIA-E in our sample was higher than would be expected (typically 5-20%). The number of 29, i.e. 39% of the total patient group, was chosen as a compromise between approximating the expected frequency of ARIA-E in anti-amyloid-beta (Aβ) antibody studies and maximization of the number of ARIA-E cases for validation.

One more possible limitation is the fact that the ARIA-E scale has not been evaluated with regard to~~for~~ its~~their~~ clinical relevance. In the current data-set many cases were only detected retrospectively and discussion about whether or not ARIA-E should be avoided are ongoing. One of the problems is defining a clinical threshold for ARIA-E has been the lack of a good rating scale. Currently, the ARIA-E rating scale has been used in at least 3 large phase-3 clinical trials of amyloid-lowering therapies. Similarly, the online training module could easily be made available for other neuroradiologists to ensure consistency in applying the ARIA-E scale.

In conclusion, the ARIA-E rating scale is a simple and robust visual rating scale that allows for determination of ARIA-E severity and regional categorization of the various manifestations. Since the amount of research in Aβ lowering immunotherapy will likely increase, the need for a rating instrument is apparent. We demonstrated that the MRI rating scale for ARIA-E is both valid and reproducible. Its application may improve the early identification ARIA-E. It can be completed in a relatively short period of time and may be used for standardized assessment to monitor medical therapy and its management in research studies.

1. Selkoe DJ, Schenk D. Alzheimer's disease: molecular understanding predicts amyloid-based therapeutics. Annu Rev Pharmacol Toxicol 2003;43:545-84.
2. Bohrmann B, Baumann K, Benz J, et al. Gantenerumab: a novel human anti-Abeta antibody demonstrates sustained cerebral amyloid-beta binding and elicits cell-mediated removal of human amyloid-beta. J Alzheimers Dis 2012;28:49-69.
3. Wattjes MP. Structural MRI. Int Psychogeriatr 2011;23 Suppl 2:S13-24.
4. Wattjes MP, Henneman WJ, van der Flier WM, et al. ~~de Vries O, Traber F, Geurts JJ, Scheltens P, Vrenken H, Barkhof F.~~ Diagnostic imaging of patients in a memory clinic: comparison of MR imaging and 64-detector row CT. Radiology 2009;253:174-83.
5. Verdile G, Fuller S, Atwood CS, Laws SM, Gandy SE, Martins RN. The role of beta amyloid in Alzheimer's disease: still a cause of everything or the only one who got caught? Pharmacol Res 2004;50:397-409.
6. Moreth J, Mavoungou C, Schindowski K. Passive anti-amyloid immunotherapy in Alzheimer's disease: What are the most promising targets? Immun Ageing 2013;10:18.
7. Roher AE, Cribbs DH, Kim RC, et al. Bapineuzumab alters abeta composition: implications for the amyloid cascade hypothesis and anti-amyloid immunotherapy. PLoS One 2013;8:e59735.
8. Bard F, Cannon C, Barbour R, et al. Peripherally administered antibodies against amyloid beta-peptide enter the central nervous system and reduce pathology in a mouse model of Alzheimer disease. Nat Med 2000;6:916-9.
9. Sperling RA, Jack CR, Jr., Black SE, et al~~Frosch MP, Greenberg SM, Hyman BT, Scheltens P, Carrillo MC, Thies W, Bednar MM, Black RS, Brashear HR, Grundman M, Siemers ER, Feldman HH, Schindler RJ~~. Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: recommendations from the Alzheimer's Association Research Roundtable Workgroup. Alzheimers Dement 2011;7:367-85.
10. Barakos J, Sperling R, Salloway S, et al. MR imaging features of amyloid-related imaging abnormalities. AJNR Am J Neuroradiol 2013;34:1958-65.
11. Salloway S, Sperling R, Gilman S, et al. A phase 2 multiple ascending dose trial of bapineuzumab in mild to moderate Alzheimer disease. Neurology 2009;73:2061-70.
12. Carlson C, Estergard W, Oh J, et al~~Suhy J, Jack CR, Jr., Siemers E, Barakos J~~. Prevalence of asymptomatic vasogenic edema in pretreatment Alzheimer's disease study cohorts from phase 3 trials of semagacestat and solanezumab. Alzheimers Dement 2011;7:396-401.
13. Sperling R, Salloway S, Brooks DJ, et al. Amyloid-related imaging abnormalities in patients with Alzheimer's disease treated with bapineuzumab: a retrospective analysis. Lancet Neurol 2012;11:241-9.
14. Wahlund LO, Barkhof F, Fazekas F, et al. European Task Force on Age-Related White Matter C. A new rating scale for age-related white matter changes applicable to MRI and CT. Stroke 2001;32:1318-22.
15. Barkhof F, Daams M, Scheltens P, et al. ~~Brashear HR, Arrighi HM, Bechten A, Morris K, McGovern M, Wattjes MP~~. An MRI rating scale for amyloid-related imaging abnormalities with edema or effusion. AJNR Am J Neuroradiol 2013;34:1550-5.
16. Black RS, Sperling RA, Safirstein B, et al~~Motter RN, Pallay A, Nichols A, Grundman M~~. A single ascending dose study of bapineuzumab in patients with Alzheimer disease. Alzheimer Dis Assoc Disord 2010;24:198-203.

17. Boche D, Zotova E, Weller RO, et al~~Love S, Neal JW, Pickering RM, Wilkinson D, Holmes C, Nicoll JA~~. Consequence of Abeta immunization on the vasculature of human Alzheimer's disease brain. Brain 2008;131:3299-310.

Table 1. Characteristics of the study population.

	ARIA-E	No ARIA-E	Total
Number of subjects	29	46	75
Age (mean)(±SD)	67.4 (8.4)	67.2 (8.4)	67.3 (8.4)
MMSE (mean)(±SD)	20.6 (3.0)	20.9 (2.8)	20.8 (2.9)
DAD (mean)(±SD)	85.4 (15.4)	86.8 (14.0)	86.2 (14.5)
Female, n (%)	20 (68.9%)	26 (57.7%)	46 (62.1%)
ApoE e4 Carriers, n (%)	22 (78.5%)	34 (77.2%)	56 (77.7%)
ApoE e4 homozygotes	7 (25.0%)	12 (27.2%)	19 (26.3%)
Bapineuzumab			
0.15 mg/kg, n (%)	3 (10.3%)	10 (22.2%)	13 (17.5%)
0.5 mg/kg, n (%)	3 (10.3%)	14 (31.1%)	17 (22.9%)
1.0 mg/kg, n (%)	10 (34.4%)	13 (28.8%)	23 (31.0%)
2.0 mg/kg, n (%)	13 (44.8%)	8 (17.7%)	21 (28.3%)

Values are means ±SD (standard deviation). ARIA-E (Amyloid-related imaging abnormalities with edema or effusion). MMSE: Mini-Mental State Examination. DAD: Disability Assessment for Dementia, APOE: apolipoprotein E. n: number of patients.

Table 2. Overview of the ARIA-E Rating Scale

Abnormality	Parenchymal Hyperintensity
	Sulcal Hyperintensity
	Swelling
Largest cross-sectional diameter	Score 0 – none
	Score 1 - monofocal ≤2 cm
	Score 2 – multifocal ≤2cm
	Score 3 – any lesion>2 but <4cm
	Score 4 – any lesion ≥ 4 cm
	Score 5 - entire lobe
Side and region	Right/Left
	Frontal
	Parietal
	Occipital
	Temporal
	Central
	Infratentorial

ARIA-E: Amyloid-related imaging abnormalities with edema or effusion

Table 3. Median of ARIA-E score, parenchymal hyperintensity score, sulcal hyperintensity score and swelling score by rater in ARIA-E cases

Rater	1		2		3		4		5	
	M	R	M	R	M	R	M	R	M	R
ARIA-E	12	(0-50)	7	(0-40)	12	(1-48)	5	(0-40)	8	(0-44)
PH	3	(0-32)	0	(0-13)	1	(0-24)	0	(0-14)	0	(0-17)
SH	3	(0-50)	4	(0-40)	7	(0-48)	4	(0-44)	3	(0-45)
SW	9	(0-50)	7	(0-37)	8	(0-48)	4	(0-44)	0	(0-26)

M: Median. R: Range, ARIA-E: amyloid related imaging abnormality with edema or effusion, PH: parenchymal hyperintensity, SH: sulcal hyperintensity, SW: gyral swelling

Table 4. Sensitivity and specificity (95% confidence intervals) for patients with and without ARIA-E by rater

Rater	1	2	3	4	5
Sensitivity	96.6% (0.80-1.00)	93.1% (0.76-0.99)	100% (0.85-1.00)	93.1% (0.75-0.99)	96.6% (0.80-1.00)
Specificity	97.8% (0.87-1.00)	95.7% (0.84-0.99)	100% (0.90-1.00)	100% (0.90-1.00)	100% (0.90-1.00)

ARIA-E (Amyloid-related imaging abnormalities with edema or effusion)

Table 5. Interobserver agreement Intraclass correlation coefficients (ICC) and the 95 % confidence intervals (CI) among 5 raters

MRI Findings	ICC	95% CI
ARIA-E (Parenchymal and sulcal hyperintensity)	0.878	0.790 - 0.936
Sulcal hyperintensity	0.915	0.854 - 0.955
Parenchymal hyperintensity	0.678	0.512 - 0.814
Swelling	0.663	0.480 - 0.807

ARIA-E (Amyloid-related imaging abnormalities with edema or effusion)

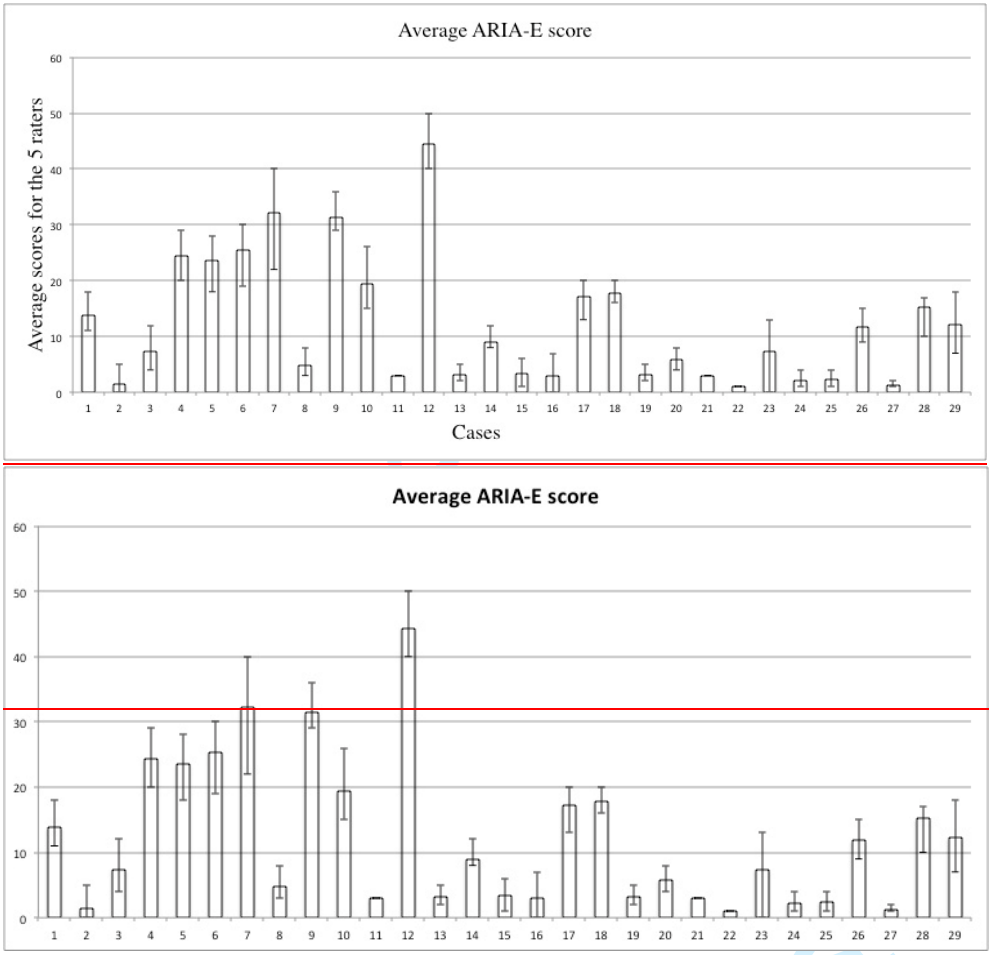

Fig 1. Average ARIA-E (Amyloid-related imaging abnormalities with edema or effusion) scores for the 5 raters in abnormal cases. Bars represent the average scores of the five raters of the maximum scores for parenchymal, sulcal hyperintensity and gyral swelling per patient of each rater summed across the 12 anatomic regions. The error bars represent the ranges of the maximum and minimum score between the 5 raters, showing that the overall variation in scores is relatively small between raters for the sum score.

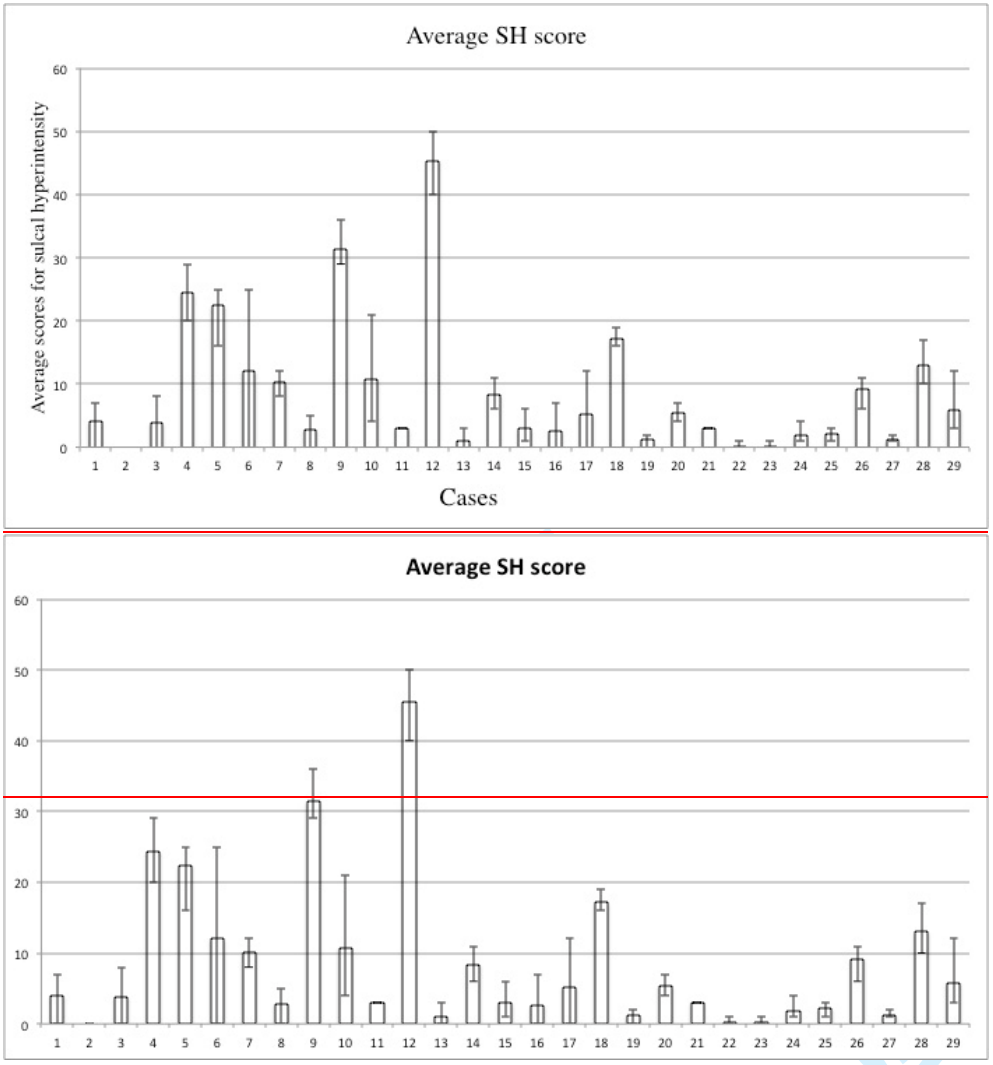

Fig 2. Distribution of Sulcal hyperintensity (SH) scores for the 5 raters in the ~~ARIA-E~~ Amyloid-related imaging abnormalities with edema or effusion cases. Bars represent average scores of the five raters per patient summed across the 12 anatomic regions. The error bars represent the ranges of the maximum and minimum score between the 5 raters.

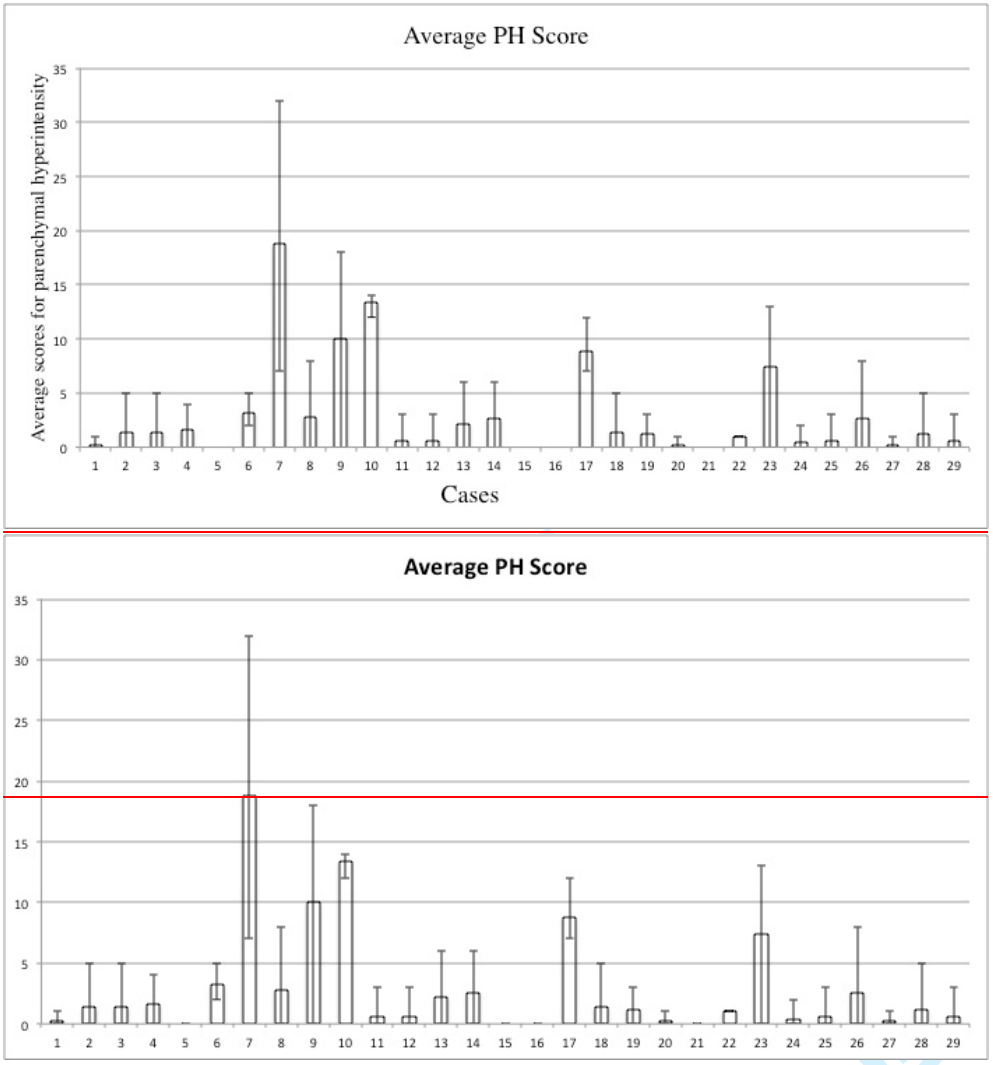

Fig 3. Distribution of Parenchymal hyperintensity (PH) scores for the 5 raters in the ARIA-E (Amyloid-related imaging abnormalities with edema or effusion) cases. Bars represent average scores of the five raters per patient summed across the 12 anatomic regions. The error bars represent the ranges of the maximum and minimum score between the 5 raters, showing that the overall variation in scores is larger than in average ARIA-E and sulcal hyperintensity ~~(SH)~~ scores between raters, indicating the higher level of complexity in detection of the PH.

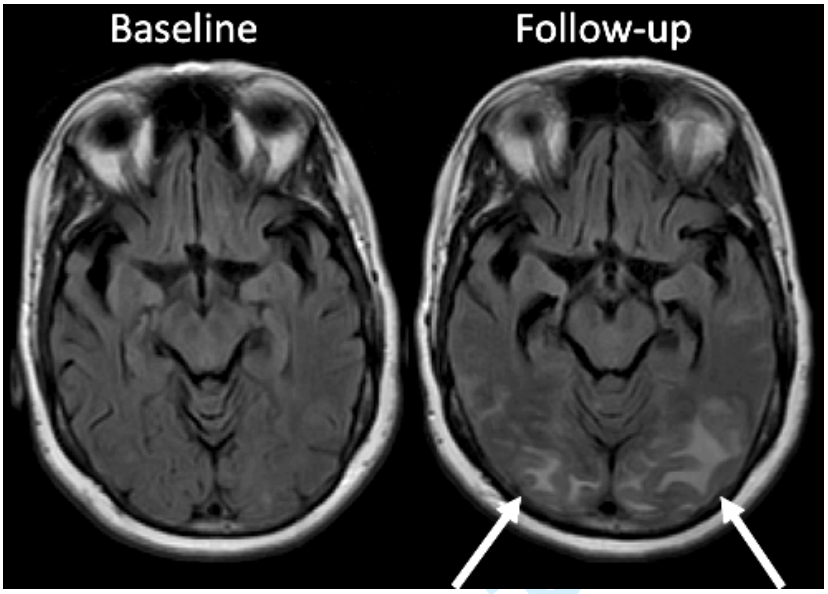

Fig 4. MR images illustrate findings leading to differences in the scoring between rater 1 and 4 for Patient 9 with widespread Amyloid-related imaging abnormalities with edema or effusion involving both occipital lobes. FLAIR (Fluid attenuation inversion recovery) images at the baseline (left side) and FLAIR images at follow up (right side) shows an extensive parenchymal hyperintensity (PH) in both occipital lobes. Rater 1 scored PH as 5 in both occipital regions and rater 4 gave a PH score of 3 for the same regions.

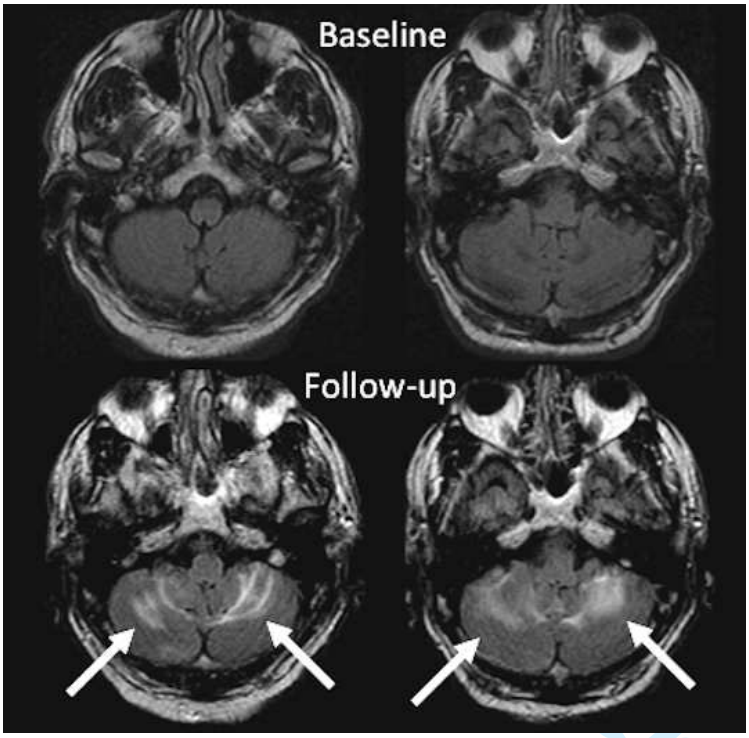

Fig 5. MR images illustrate the cause of differences in the scoring between rater 1 and 2 for patient 7 with widespread Amyloid-related imaging abnormalities with edema or effusion involving both infratentorial regions. Discrepant readings for parenchymal hyperintensity (PH) in this patient are due to difficulty distinguishing parenchymal from sulcal hyperintensity (SH). The top row shows FLAIR (Fluid attenuation inversion recovery) images at the baseline. The lower row shows FLAIR images at the follow up with PH and SH in both cerebellar hemispheres. Rater 1 gave a score of 5 for PH and SH for each side (left and right) infratentorial, while rater 2 scored 0 for PH and 5 for SH at the same regions

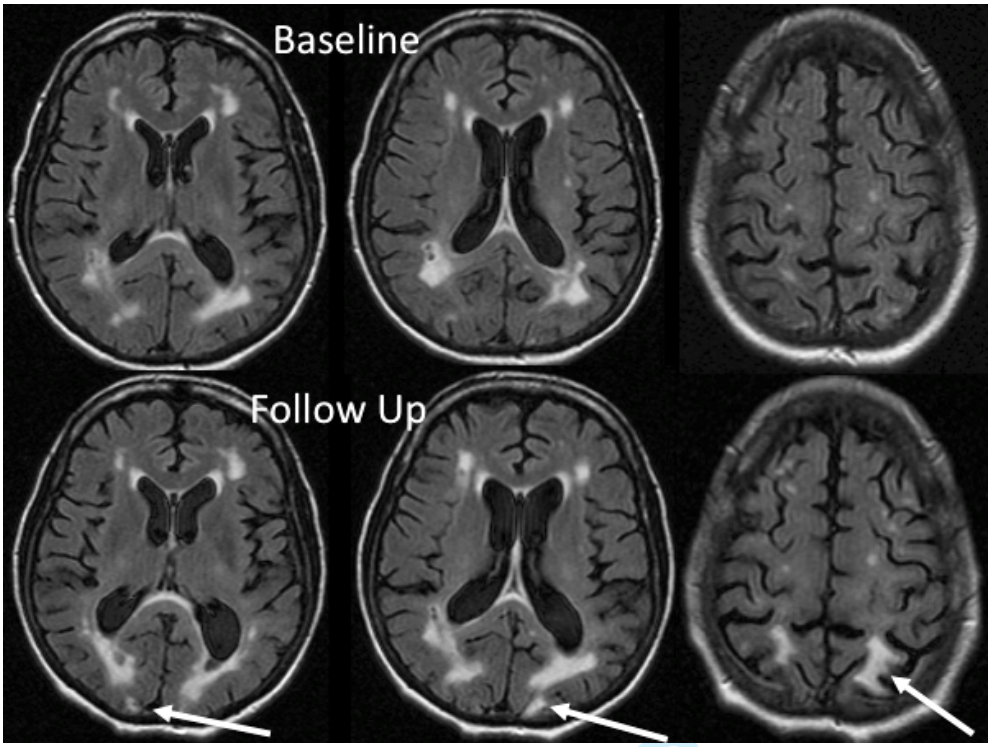

Fig 6. Patient number 22 from the ARIA-E (Amyloid-related imaging abnormalities with edema or effusion) series was not recognized by only one of the raters (rater 1) This ARIA-E patient has substantial white matter hyperintensities ~~(WMH)~~ which may visually mask an evolving parenchymal hyperintensities (PH). The top row shows FLAIR (Fluid attenuation inversion recovery) images at baseline with considerable vascular white matter lesions. The lower row shows follow up FLAIR images with PH distributed cortically at both partieto-occipital lobes (arrows)

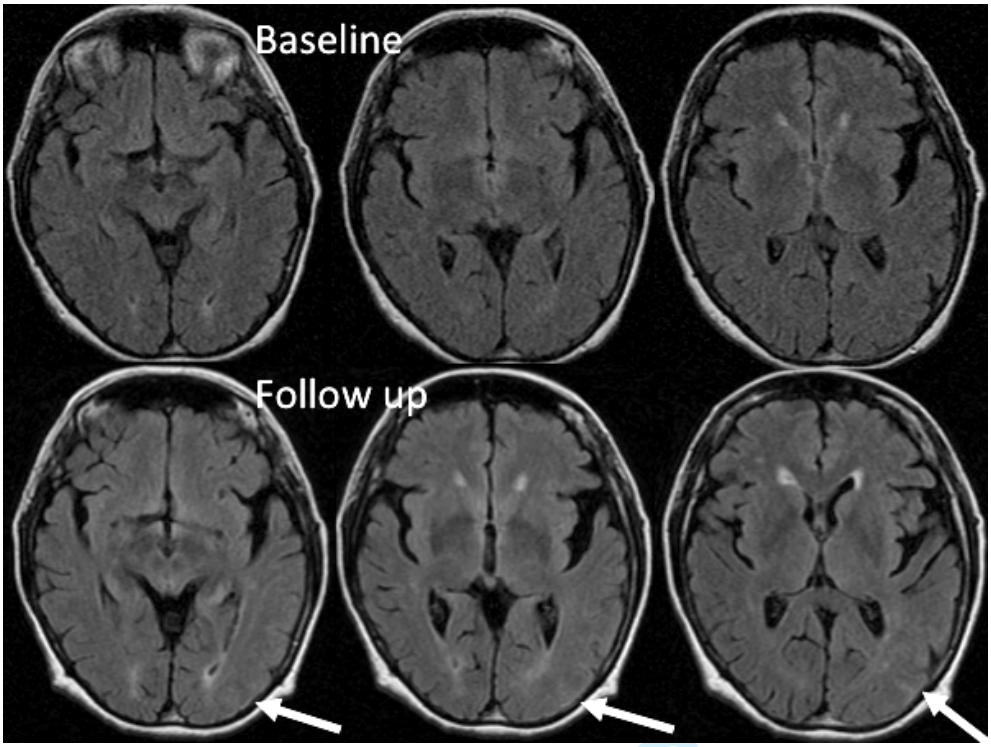

Fig 7. This ARIA-E (Amyloid-related imaging abnormalities with edema or effusion) patient presented with subtle sulcal hyperintensity (SH~~sulcal hyperintensity~~) which was recognized as due to ARIA-E based on parenchymal hyperintensity by three of the raters (rater 1, 3 and 5), but only 1 of the raters scored for SH and swelling (SW) in the left occipital lobe. The top row shows FLAIR (Fluid attenuation inversion recovery) images at baseline. The lower row shows FLAIR images at the follow up illustrating SH and SW in the left occipital lobe (arrows).