

RESEARCH ARTICLE

Comparing ARIA-E severity scales and effects of treatment management thresholds

Gregory Klein¹ | Marzia A. Scelsi² | Jerome Barakos³ | Jochen B. Fiebach⁴ |
 Luc Bracoud⁵ | Joyce Suhy⁶ | Paul Delmar¹ | Marco Lyons² | Jakub Wojtowicz¹ |
 Szofia Bullain¹ | Frederik Barkhof^{7,8} | Derk Purcell³

¹F. Hoffmann-La Roche Ltd, Basel, Switzerland

²Roche Products Ltd., Welwyn Garden City, UK

³California Pacific Medical Center, San Francisco, California, USA

⁴Center for Stroke Research Berlin, Charité Universitätsmedizin Berlin, Berlin, Germany

⁵Clario, Inc. (formerly Bioclinica, Inc.), Lyon, France

⁶Clario, Inc. (formerly Bioclinica, Inc.), San Mateo, California, USA

⁷Department of Radiology & Nuclear Medicine, Amsterdam UMC, Vrije Universiteit, Amsterdam, Netherlands

⁸Queen Square Institute of Neurology and Centre for Medical Image Computing, University College London, London, UK

Correspondence

Gregory Klein, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 170, 4070 Basel, Switzerland.

E-mail: gregory.klein@roche.com

Abstract

Introduction: Amyloid-related imaging abnormalities–edema (ARIA-E) is associated with anti-amyloid beta monoclonal antibody treatment. ARIA-E severity may be assessed using the Barkhof Grand Total Scale (BGTS) or the 3- or 5-point Severity Scales of ARIA-E (SSAE-3/SSAE-5). We assessed inter- and intra-reader correlations between SSAE-3/5 and BGTS.

Methods: Magnetic resonance imaging scans were collected from 75 participants in the SCarlet RoAD and Marguerite RoAD studies. Three neuroradiologists reviewed scans at baseline and at follow-up. Concordance in dichotomized ARIA-E ratings was assessed for a range of BGTS thresholds.

Results: SSAE-3/5 scores correlated with BGTS scores, with high inter-reader intraclass correlation coefficients across all scales. There was high agreement in dichotomized ratings for SSAE-3 > 1 versus BGTS > 3 for all readers (accuracy 0.85–0.93) and between pairs of readers.

Discussion: SSAE-3/5 showed high degrees of correlation with BGTS, potentially allowing seamless transition from the BGTS to SSAE-3/5 for ARIA-E management.

KEYWORDS

Alzheimer's disease, amyloid-related imaging abnormalities, ARIA-E, Barkhof Grand Total Scale, correlation, edema, magnetic resonance imaging, Severity Scale of ARIA-E, SSAE

1 | BACKGROUND

Alzheimer's disease (AD) is a chronic, progressive neurodegenerative disease that causes dementia.¹ It is characterized by the neurotoxic accumulation of amyloid beta ($A\beta$) plaques comprising $A\beta$ peptides and intracellular neurofibrillary tangles containing tau protein, in the brain.^{2,3} Deposition of $A\beta$ in the parenchyma likely occurs decades before clinical symptoms manifest.²

Anti- $A\beta$ monoclonal antibodies (mAbs) represent a major area of drug development for AD and have shown promise in slowing progression, positively impacting cognitive impairment, and reducing amyloid pathology through microglial stimulation or prevention of $A\beta$ aggregation.^{4–6} Aducanumab, an anti- $A\beta$ mAb, was granted accelerated approval by the US Food and Drug Administration in 2021, while several other anti- $A\beta$ treatments, including gantenerumab, donanemab, and lecanemab, are in clinical development, with data readouts from ongoing clinical trials expected in 2022 and 2023.^{5,7,8}

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* published by Wiley Periodicals, LLC on behalf of Alzheimer's Association.

RESEARCH IN CONTEXT

- 1. Systematic review:** The authors used PubMed to identify literature relating to amyloid-related imaging abnormalities–edema (ARIA-E) severity scales and studies of anti-amyloid beta monoclonal antibodies in which ARIA-E radiographic severity has been assessed using severity scales.
- 2. Interpretation:** High levels of correlation were found between the Barkhof Grand Total Scale (BGTS) and the simpler 3- and 5-point Severity Scales of ARIA-E (SSAE-3/5), suggesting the potential for translation between treatment management rules using these scales.
- 3. Future directions:** The greater granularity of SSAE-5 compared with SSAE-3 could allow more patients to benefit from uninterrupted treatment. This hypothesis is being explored in the ongoing Post-GRADUATE open-label extension study of gantenerumab.

HIGHLIGHTS

- A simple rating scale is needed to rate severity of amyloid-related imaging abnormalities–edema (ARIA-E) in the clinical setting.
- The 3- and 5-point Severity Scales of ARIA-E (SSAE-3/5) and Barkhof Grand Total Scale (BGTS) have good correlation and agreement between readers.
- Treatment management thresholds show high concordance of decision between scales.
- SSAE-3/5 may be more practical for assessing ARIA-E than the BGTS in some settings.
- Granularity of SSAE-5 versus -3 could allow more flexibility in clinical interventions.

A side effect associated with anti-A β therapies is amyloid-related imaging abnormalities (ARIA), which appear either as vasogenic edema in the parenchyma or sulcal effusions in leptomeninges (ARIA-E), or as microhemorrhages in the parenchyma or superficial siderosis in leptomeninges (ARIA-H) on magnetic resonance imaging (MRI) scans.⁹ ARIA-like findings have also been identified in other disease states such as cerebral amyloid angiopathy and cerebral amyloid angiopathy-related inflammation and can also occur spontaneously.^{10,11} Anti-A β -treatment-induced ARIA is a commonality of mAbs targeting fibrillar amyloid; it is mostly clinically asymptomatic, and as such has typically been identified through routine MRI scans during clinical trials; therefore, routine monitoring for ARIA in patients receiving anti-A β therapies is recommended.^{7,9} ARIA-E is best visualized via a T2-fluid-attenuated inversion recovery (T2-FLAIR) sequence.⁹ Depending

TABLE 1 SSAE definition

ARIA-E extent	ARIA-E focality	SSAE-3	SSAE-5
No ARIA-E	N/A	0	0
< 5 cm	Monofocal	1 (mild)	1 (mild)
	Multifocal	2 (moderate)	2 (mild+)
5 to 10 cm	Monofocal		3 (moderate)
	Multifocal		4 (moderate+)
> 10 cm	Monofocal	3 (severe)	5 (severe)
	Multifocal		

Abbreviations: ARIA-E, amyloid-related imaging abnormalities–edema; N/A, not applicable; SSAE, Severity Scale of ARIA-E.

on the radiographic severity and symptomatology of ARIA-E, treatment may continue, be modified, or suspended. Quantitative, objective ARIA-E scales are required to define severity cut-offs to guide trial investigators and prescribers in clinical practice on whether dose adjustment or suspension is required.

Scales used to quantify ARIA-E radiographic severity in clinical research include the Barkhof Grand Total Scale (BGTS),¹² and the 3- and 5-point Severity Scales of ARIA-E (SSAE-3, SSAE-5).¹³ The BGTS is a comprehensive 60-point rating system that assesses ARIA-E in six regions of the brain (frontal, parietal, occipital, temporal, central, and infratentorial) on both the left and right sides, yielding 12 separate anatomic locations for evaluation (details in Table S1 in supporting information).¹²

By contrast, the SSAE-3 and SSAE-5 are simpler rating systems to assess ARIA-E severity (Table 1) based upon a single linear measurement of the largest area of lesion. Both SSAE scales define severity as the result of measured spatial ARIA-E extent and distribution. Extent is defined as the single greatest measured diameter of the ARIA-E lesions, classified as < 5 cm, 5 to 10 cm, or > 10 cm, and their multiplicity. In assessing extent, no distinction is made between parenchymal hyperintensities, sulcal hyperintensities, and sulcal effacement/swelling (Table 1).¹³ Spatial distribution is captured as the number of non-contiguous regions of ARIA-E, thus determining whether the ARIA-E is mono- or multifocal in nature. The SSAE-3 classifies ARIA-E on a scale of 0 to 3 as 0 = no ARIA-E, 1 = mild (T2-FLAIR hyperintensity confined to sulcus and or cortex/subcortical white matter in one location and < 5 cm), 2 = moderate (T2-FLAIR hyperintensity 5 to 10 cm, or more than one site of involvement, each measuring \leq 10 cm), or 3 = severe (T2-FLAIR hyperintensity measuring > 10 cm, one or more separate sites of involvement may be noted).¹³ The SSAE-5 adds two additional severity ratings, providing information related to multiplicity of lesions: mild and moderate denote monofocal ARIA-E, whereas mild+ and moderate+ denote multifocal ARIA-E. Scores of 0 and 1 on both SSAE-3 and SSAE-5 are identical, as are scores of SSAE-3 = 3 and SSAE-5 = 5 (Table 1). SSAE-5 gives greater granularity than SSAE-3 between SSAE-5 scores 2 to 4, which allows for more precise treatment management interventions compared with the moderate rating of the SSAE-3.

The BGTS was used to assess ARIA-E in two Phase III studies of gantenerumab: SCarlet RoAD (SR; NCT01224106)^{4,14} and Marguerite RoAD (MR; NCT02051608).¹⁵ Both studies evaluated the efficacy (as measured by cognition and function), safety, and pharmacokinetics of subcutaneous (SC) gantenerumab. SR included patients with mild cognitive impairment due to AD, and MR included patients with mild AD dementia. In the double-blind parts of both studies, study participants were randomized to receive lower-level doses of SC gantenerumab (105 mg or 225 mg every 4 weeks [Q4W] in SR, 105 mg Q4W for 6 months followed by 225 mg Q4W in MR), or placebo. In the open-label extension (OLE) parts of both studies, participants who completed the double-blind part of the trial received SC gantenerumab at doses up to 1200 mg Q4W for 3 additional years. The BGTS is also used in the ongoing Phase III GRADUATE trials, which are testing a higher dose of gantenerumab (given as 510 mg every 2 weeks; GRADUATE 1, NCT03444870; GRADUATE 2, NCT03443973). ARIA-E management rules in these two gantenerumab studies mandated treatment interruption for any ARIA-E radiographic severity that was symptomatic, or for asymptomatic ARIA-E with BGTS > 3 until ARIA-E resolution, whereas BGTS ≤ 3 allowed for continued dosing at the same level until ARIA-E resolution (Figure S1 in supporting information).

The SSAE-3 assessing ARIA-E severity has been used in most AD clinical trials conducted to date, including those of solanezumab, donanemab, lecanemab, and aducanumab,¹⁶⁻¹⁹ and is specified in the prescribing information for aducanumab.⁷ In the EMERGE (NCT02484547) and ENGAGE (NCT02477800) Phase III double-blind and randomized clinical trials of aducanumab, patients were randomized to receive a low or high dose of intravenous aducanumab, or placebo, and the SSAE-3 was used to measure radiographic severity of ARIA-E.²⁰ Asymptomatic patients with mild radiographic ARIA-E (SSAE-3 = 1) could continue dosing at the same dosing schedule, but patients with asymptomatic moderate or severe radiographic ARIA-E (SSAE-3 > 1) had dosing suspended (Figure S1).¹⁷ Using the SSAE-3 and SSAE-5 in clinical practice would allow for easier ARIA-E characterization. The Post-GRADUATE (NCT04374253)²¹ and SKYLINE (NCT05256134)²² ongoing clinical trials of gantenerumab are currently using the SSAE-5 for the radiographic severity assessment of ARIA-E, and this scale is potentially more clinically valuable than the SSAE-3 due to a greater range of dosing intervention options available.

Validation analyses showed that both the BGTS and the SSAE-3 have high intra- and inter-rater agreement.¹³ Because these scales have been used independently in clinical trials, being able to show treatment decision translatability among the BGTS, SSAE-3, and SSAE-5 scores would be useful for comparing different datasets, and in the future, for harmonizing ARIA-E management recommendations. We previously used the median score across three readers and its interquartile range to align BGTS scores with SSAE-3 and SSAE-5.²³ Here we assess the inter- and intra-reader correlation among the SSAE-3, SSAE-5, and BGTS.

2 | METHODS

2.1 | Study design

MRI scans were collected from a sample of 75 participants, from the double-blind and OLE phases of SR and MR, with mild cognitive impairment due to AD (SR), or mild AD dementia (MR; 70 with previously detected ARIA-E, 5 without ARIA-E from original BGTS assessment). Participants without ARIA-E were included to provide blinding for the readers. Only the baseline MRI scan and a second scan taken after the incident detection of ARIA-E (for those cases included in the analysis with ARIA-E) were used for this study. The sample was chosen to be representative of the original BGTS distribution seen in SR and MR double-blind and OLE study participants, and is enriched for ARIA-E; therefore, it is not reflective of the gantenerumab ARIA-E incidence rate (30.5% and 31.9% in the SR and MR OLEs, respectively [Roche data on file]). ARIA-E severity was assessed using the SSAE-3, SSAE-5, and BGTS by three neuroradiologists experienced with detection of ARIA-E, who were blinded to prior read results. Assessment was performed independently by the three readers, with no discussion or consensus in case of disagreement. The readers were provided with paired T2-FLAIR and T2*-gradient echo MRI scans (5 mm axial slice, no gaps, 256 × 256 matrix) taken at baseline and at a follow-up visit, for the 70 study participants for whom ARIA-E was reported during on-study monitoring, and for the 5 study participants with no ARIA-E reported.

2.2 | Treatment management threshold

ARIA-E ratings were dichotomized (BGTS/SSAE-3 higher than a certain threshold) and the concordance in these dichotomized ratings was assessed quantitatively for a range of BGTS thresholds.

Leveraging the availability of both BGTS and SSAE-3 ratings for each scan by three experienced readers, and looking for the optimal BGTS severity cut-off to match a cut-off of > 1 on the SSAE-3 and SSAE-5 scales, comparisons were made: between readers (i.e., how many times two readers rating the same scan would give the same dichotomized rating by using either the BGTS or SSAE-3); and within reader (i.e., how many times a given reader would give the same dichotomized rating by using the BGTS and SSAE-3). Because scores of 0 and 1 are identical for both SSAE-3 and SSAE-5, the dichotomization threshold is the same for both scales using the cut-offs of SSAE-3/SSAE-5 > 1. Comparison of dichotomized ratings of BGTS scores with SSAE-3 scores can therefore be considered equivalent to BGTS scores versus SSAE-5 scores.

2.3 | Statistical analysis

Agreement between and within readers was quantified for a range of BGTS thresholds by computing confusion matrices of dichotomized rating (score higher than a certain threshold: yes/no), then deriving

accuracy (95% confidence interval [CI]) as the ratio of the trace of the confusion matrix to the sum of all matrix elements. Cohen's kappa statistic was also calculated from the same confusion matrix. The primary analysis quantified the treatment management agreement for the BGTS and SSAE thresholds used in past studies (BGTS > 3, SSAE-3 > 1), but a receiving operator characteristic analysis was also performed to consider agreement using other possible BGTS thresholds against the SSAE > 1 threshold. Additionally, we assessed agreement between BGTS and SSAE ratings (both 3-point and 5-point) for each reader separately using Spearman's rank correlation coefficient and Cohen's kappa statistic.

3 | RESULTS

3.1 | Baseline characteristics

Baseline demographic characteristics of the 75 participants were generally well balanced across the SR and MR groups (Table S2 in supporting information). Of the 75 participants, the mean (standard deviation [SD]) age was 70.2 (7.8) years; 54.7% were female; 88.0% were White. In the SR cases, 21 participants (75.0%) were apolipoprotein E (APOE) ϵ 4 allele carriers, compared with 33 (70.2%) in the MR cases; overall, 54/75 (72.0%) participants in this analysis were APOE ϵ 4 allele carriers. Thirteen participants (17.3%) had ARIA-H at baseline.

3.2 | Descriptive summaries of ARIA rating scales

The mean (SD) BGTS severity scores for the readers 1, 2, and 3, respectively, were: 5.92 (5.94), 4.48 (4.69), and 6.24 (6.42). The BGTS, SSAE-3, and SSAE-5 had high inter-rater intra-class correlation (ICC) values: inter-rater ICCs (lower bound, upper bound) for the BGTS, SSAE-3, and SSAE-5 were 0.88 (0.84–0.91), 0.87 (0.83–0.91), and 0.86 (0.81–0.90), respectively. There was close agreement between all readers for both SSAE-3 and SSAE-5 compared with the BGTS (Figure S2 in supporting information).

Numbers and percentages of scans scored in each category for both SSAE-3 and SSAE-5 for each reader are reported in Table S3 in the supporting information. Of the lesions scored as moderate (SSAE-3 = 2), on average 57.3% of them were multifocal and less than 5 cm in extent (69.2%, 51.4%, and 51.4% for readers 1, 2, and 3, respectively), whereas on average 7.4% of lesions were unifocal and 5 to 10 cm in extent (2.6%, 8.1%, and 11.4% for readers 1, 2, and 3, respectively).

3.3 | Correlation between the SSAE-3/5 and BGTS

SSAE-3 and SSAE-5 scores correlated well with BGTS scores (Figure 1): Spearman's rank correlations (95% CIs) between the SSAE-3 and BGTS were 0.87 (0.80–0.91), 0.82 (0.72–0.88), and 0.87 (0.80–0.92) for readers 1, 2, and 3, respectively; Spearman's rank correlations (95% CIs) between the SSAE-5 and BGTS were 0.90 (0.84–0.93), 0.82 (0.73–0.88), 0.91 (0.87–0.94) for readers 1, 2, and 3, respectively. Comparison

of median BGTS scores to SSAE-3 and SSAE-5 scores is reported in the supporting information.

3.4 | Agreement in treatment management thresholds

As the thresholds for treatment management of SSAE-3 and SSAE-5 are identical, only the agreement in dichotomized ratings between the BGTS and SSAE-3 was analyzed. There was high agreement in dichotomized ratings for the SSAE-3 > 1 threshold versus BGTS > 3 for all three readers, with accuracies ranging from 0.85 to 0.93 (Table 2). The highest concordance in dichotomized ratings based on SSAE-3 > 1 and BGTS was achieved with BGTS > 3, for all three readers (Figure 2), indicating that treatment management decisions with the two rating scales would be consistent at the selected threshold. Cohen's kappa statistics were 0.87, 0.71, and 0.79 for readers 1, 2, and 3, respectively. Intra-reader SSAE-3 > 1 threshold versus BGTS > 3 dichotomized rating true positive and false negative rates were high and low, respectively (Figure S3 in supporting information).

Comparing dichotomized ratings between two readers using the same scale, there was also a high agreement, using either BGTS > 3 or SSAE-3 > 1 (Table 3). The agreement in dichotomized rating between two readers using the same scale (Table 3) is similar to the agreement in dichotomized rating between a single reader using BGTS versus SSAE-3 scales (Table 2).

Comparing dichotomized ratings based on BGTS and SSAE-3 for individual readers, accuracies (95% CI) for the SSAE-3 > 1 threshold versus BGTS > 3 were 0.93 (0.83–0.97), 0.85 (0.75–0.92), and 0.89 (0.80–0.95) for readers 1, 2, and 3, respectively (Figure 2). Cohen's kappa statistics were 0.84, 0.71, and 0.79 for readers 1, 2, and 3, respectively. Intra-reader concordance in dichotomized ratings at SSAE-3 > 1 and BGTS > 3 was similar to inter-reader concordance in dichotomized ratings at BGTS > 3 (Table 3).

As expected, there is not a perfect dichotomized rating agreement between the BGTS and SSAE-3 scales. False positive rates range between 11% and 23%, while false negative rates range between 3% and 10% for the three readers (Table 2). The frequencies of discrepant cases were 8%, 15%, and 11% for readers 1, 2, and 3, respectively.

Examples of BGTS and SSAE cases that could result in concordant and discrepant treatment management are shown in Figure 3. The top row shows five concordant cases in which the SSAE-5 score of each case resulted in a BGTS score near the mean for that corresponding distribution as seen in Figure 1. The bottom row shows four examples of discrepant cases. In the two cases on the bottom left, a potential treatment management discrepancy is caused by two distinct, but very small, areas of parenchymal hyperintensities in the left and right frontal lobes, each less than 5 cm in size, corresponding to an SSAE-5 severity of 2, but only a score of 3 and 2 on BGTS, respectively. In the third example on the bottom row, a single < 5 cm area of parenchymal hyperintensity in the right parieto-occipital lobe corresponds to an SSAE-5 severity of 1, but because it is more than 4 cm in size, leads to a BGTS score of 7. Finally, the fourth discrepant case shows a single < 5 cm area of ARIA-E in the right occipital lobe, corresponding to an SSAE

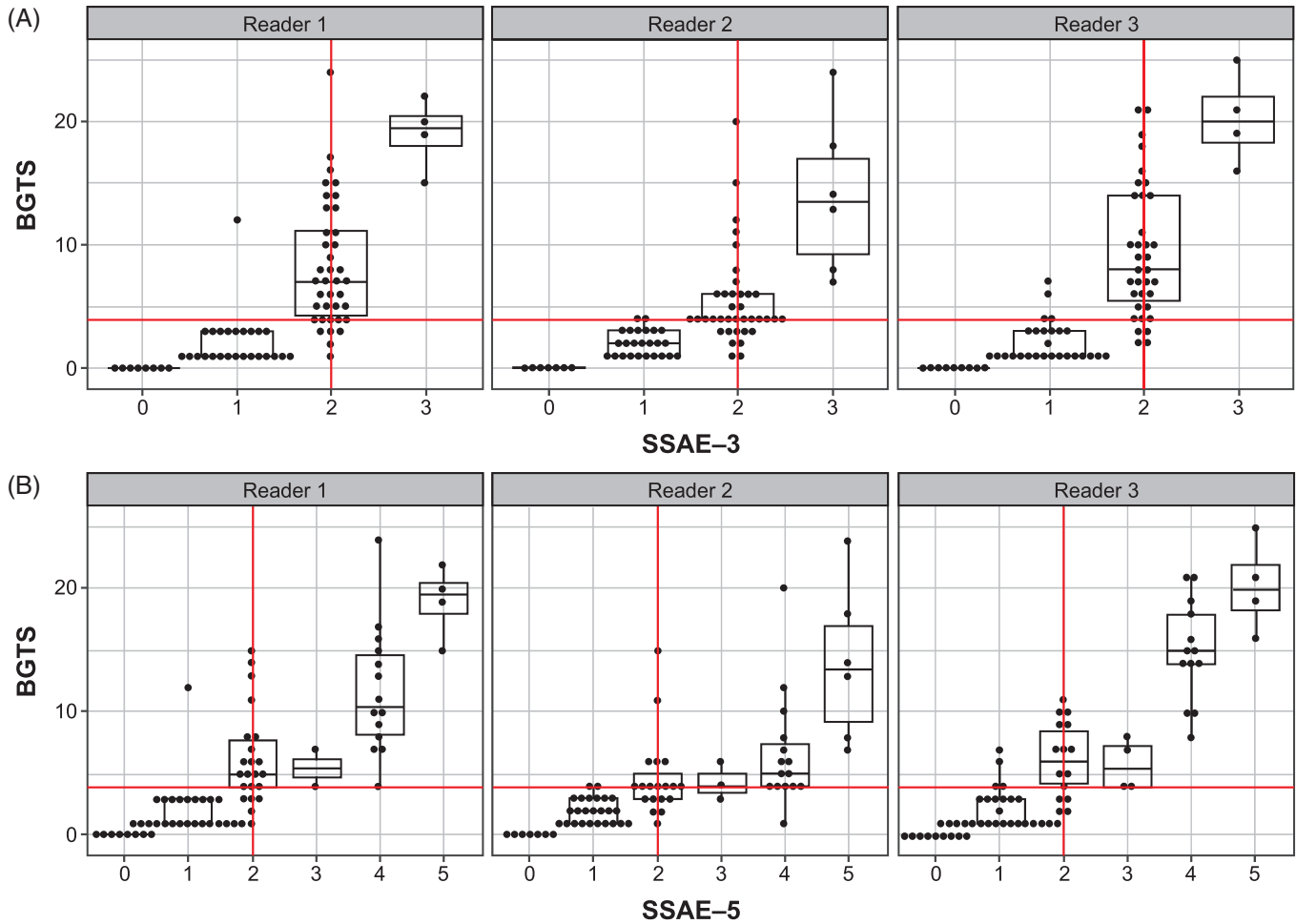


FIGURE 1 Comparison of individual readers using the BGTS and SSAE-3 (A) and using the BGTS and SSAE-5 (B). Treatment thresholds of BGTS > 3 and SSAE-3/ SSAE-5 > 1 are marked in red. ARIA-E, amyloid-related imaging abnormalities–edema; BGTS, Barkhof Grand Total Scale; SSAE, Severity Scale of ARIA-E

TABLE 2 High agreement in dichotomized ratings between BGTS > 3 and SSAE-3 > 1

SSAE-3 > 1 vs. BGTS > 3	Reader 1	Reader 2	Reader 3
Accuracy (95% CI)	0.93 (0.83–0.97)	0.85 (0.75–0.92)	0.89 (0.80–0.95)
False positive rate ^a	0.11	0.23	0.11
False negative rate ^a	0.03	0.06	0.10

Abbreviations: ARIA-E, amyloid-related imaging abnormalities–edema; BGTS, Barkhof Grand Total Scale; CI, confidence interval; SSAE, Severity Scale of ARIA-E.

^aIn the context of this analysis, a false positive means SSAE-3 ≥ 2 but BGTS < 4 (which would lead to a decision to suspend dosing according to SSAE-3 but not BGTS), whereas a false negative means SSAE-3 < 2 but BGTS > 3 (which would lead to a decision to continue dosing according to SSAE-3 but not to BGTS).

TABLE 3 High agreement in dichotomized ratings between pairs of readers using BGTS > 3 or SSAE-3 > 1

Accuracy (95% CI)	Reader 1 vs. 2	Reader 1 vs. 3	Reader 2 vs. 3
BGTS > 3	0.89 (0.80–0.95)	0.93 (0.85–0.98)	0.91 (0.82–0.96)
SSAE-3 > 1	0.95 (0.87–0.99)	0.92 (0.83–0.97)	0.92 (0.83–0.97)

Abbreviations: ARIA-E, amyloid-related imaging abnormalities–edema; BGTS, Barkhof Grand Total Scale; CI, confidence interval; SSAE, Severity Scale of ARIA-E.

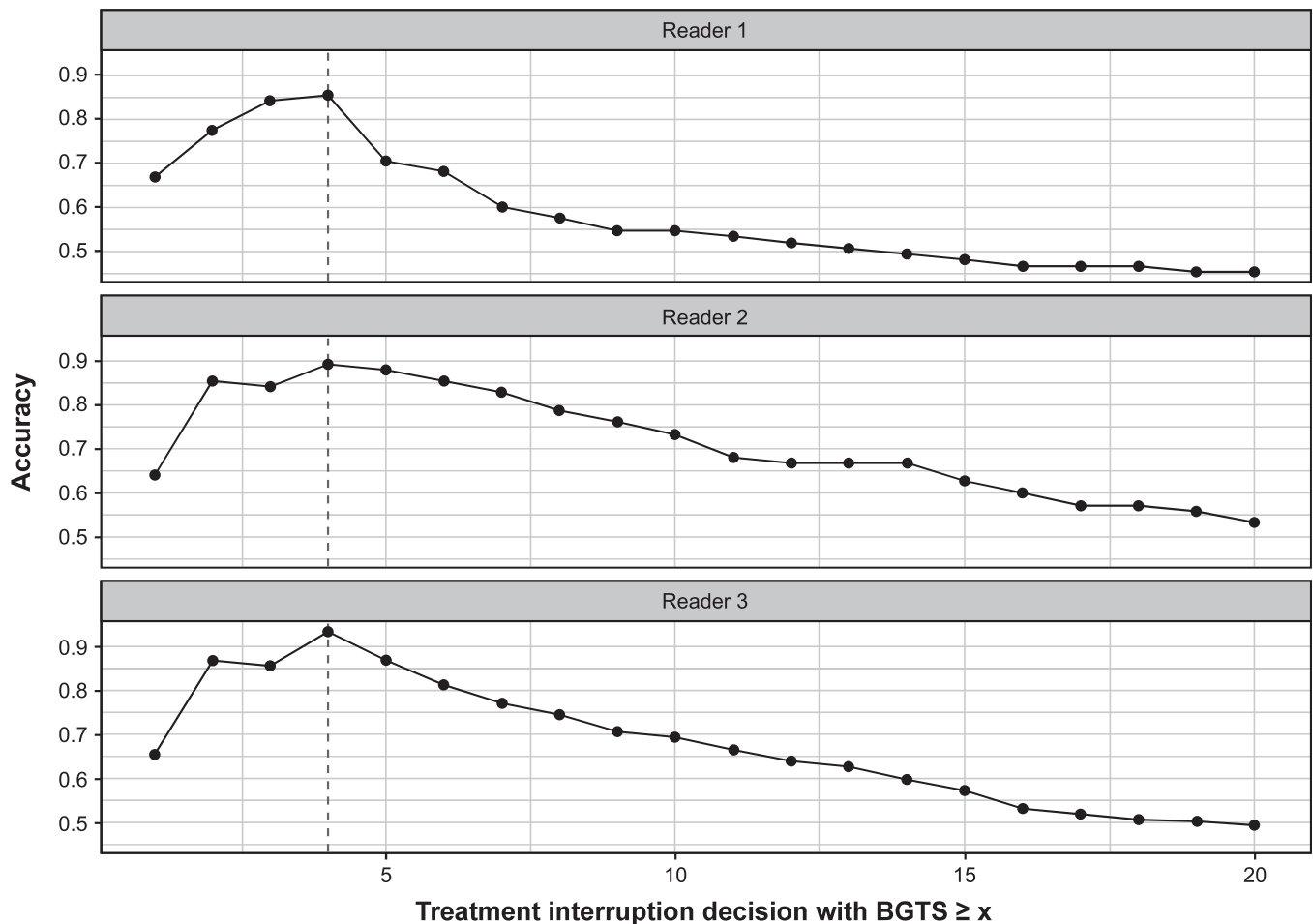


FIGURE 2 Concordance in dichotomized ratings comparing SSAE-3 > 1 threshold with possible BGTS thresholds for individual readers. ARIA-E, amyloid-related imaging abnormalities–edema; BGTS, Barkhof Grand Total Scale; SSAE, Severity Scale of ARIA-E

severity of 1, but scoring 4 on the BGTS due to the size between 2 and 4 cm.

4 | DISCUSSION

The ability to assess ARIA-E semi-quantitatively in a research setting and in clinical practice is necessary for managing treatment with anti- $A\beta$ mAbs. Although the BGTS is a valuable tool in a research setting, with the recent and future marketing approvals of $A\beta$ mAbs, a simple, practical scale for real-world clinical use is needed. During clinical trials, scans are read centrally by neuroradiologists experienced in detection and scoring of ARIA-E; however, in clinical practice, it is anticipated that reading will be performed by local neuroradiologists or general radiologists who would need to assess ARIA-E presence and severity accurately as part of their practice.^{12,24} While BGTS scoring requires that all brain locations be evaluated separately, distinguishing between, and individually scoring, sulcal hyperintensities, parenchymal hyperintensities, and swelling, the SSAE-3 and SSAE-5 scales only require a global assessment and minimal spatial characterization to calculate the final score. Although this study did not record actual time required to perform the ARIA-E assessments, which could be a topic

for future research, the readers in this study estimate that the BGTS assessment, including time required to complete a report, requires approximately three times as long as for the SSAE scales. The time required to train on the BGTS is also significantly longer than for the SSAE scales. As such, the SSAE-3 and SSAE-5 are easier to administer for practical use in a clinical setting than the BGTS, potentially allowing for a more direct characterization of ARIA-E. Furthermore, to avoid confusion among radiologists, to simplify training, and to streamline the workflow in radiology settings, it is important to implement standardized ARIA-E severity scales, such as the SSAE scales, that could be used similarly for all anti- $A\beta$ mAb therapies.

This study has shown that the SSAE-3 and SSAE-5 scales are well correlated with the BGTS, with concordance on a treatment management threshold of SSAE-3 > 1 corresponding to a threshold of BGTS > 3 for all readers considered. This allows for a direct and routine translation of ARIA-E management rules between scales. High rates of false negatives would be of concern for a transition from BGTS to SSAE-3 that safeguards patient safety. However, the proposed cut-off of SSAE-3 > 1 shows false positive rates higher than false negative rates consistently across readers, allowing for treatment management aligned with (and in many cases even more conservative than) BGTS > 3.

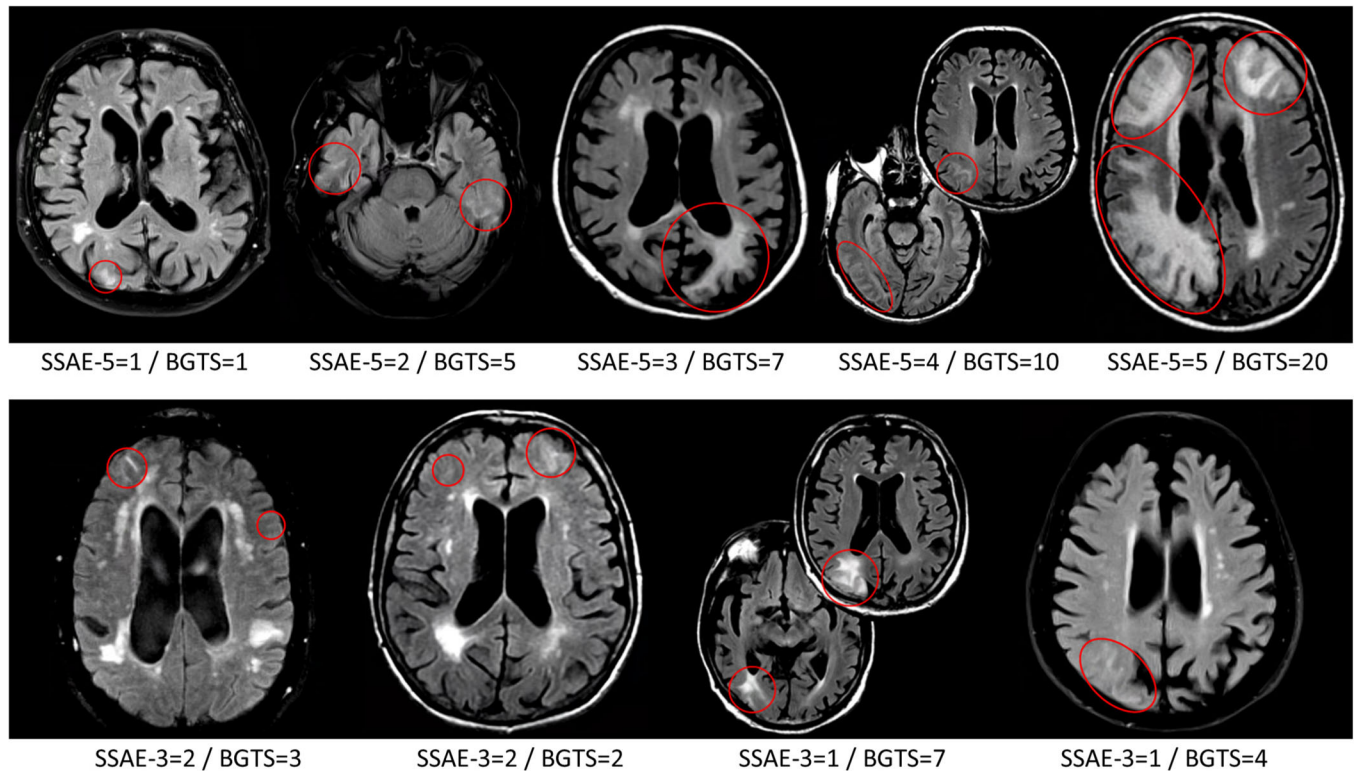


FIGURE 3 Examples where BGTS and SSAE ratings are concordant (top row) and discrepant (bottom row). BGTS, Barkhof Grand Total Scale; SSAE, Severity Scale of amyloid-related imaging abnormalities—edema

As expected, there is not perfect agreement between the BGTS and SSAE scales given the differences in extent thresholds and in lobular anatomical characterization of ARIA-E lesions. Multifocal ARIA-E occurrence would be scored as SSAE-3 = 2 (leading to treatment interruption), whereas it may be scored as ≤ 3 on the BGTS if the extent of each focus is less than 2 cm. Inversely, a monofocal ARIA-E occurrence of less than 5 cm would be scored as SSAE-3 = 1, whereas it could be > 3 on the BGTS if its size is between ≥ 4 cm and < 5 cm or if it extends over multiple adjacent anatomical locations. The separate assessment of different brain lobes in the BGTS causes a wide range of scores that correspond to SSAE-3 = 2 without improving the treatment management.

The SSAE-3 and SSAE-5 provide an alternative when making decisions regarding treatment management for patients with ARIA-E, using scales that may be easier to use compared with the BGTS in a clinical environment. Both SSAE-3 and SSAE-5 scores showed a high degree of inter-reader agreement and correlation with the BGTS, when scoring ARIA-E severity. ICC inter-reader agreement was similar using the SSAE-3 (ICC = 0.87) or SSAE-5 (ICC = 0.86). The greater granularity of the SSAE-5 compared with the SSAE-3 could allow a greater spectrum of clinical intervention. This hypothesis is being explored in the ongoing Post-GRADUATE OLE study of gantenerumab.²¹

A limitation of the study is that the rating scales have only been validated in terms of inter-rater agreement by three specialized and experienced readers. Further validation of the SSAE-3 and SSAE-5 visual rating is needed in a more heterogeneous group of raters in

terms of expertise. This is also highlighted as a limitation by Barkhof et al. evaluating the BGTS.¹² Another limitation may be that the SSAE treatment management thresholds are based around ARIA-E lesions of 5 to 10 cm, whereas the BGTS thresholds are based around lesions of 2 to 4 cm. Future work could explore potential alternative treatment management thresholds for SSAE.

5 | CONCLUSION

This study shows that both SSAE-3 and SSAE-5 are highly correlated with the BGTS; when ARIA-E ratings are dichotomized, the thresholds of BGTS > 3 , SSAE-3 > 1 , and SSAE-5 > 1 ensure maximum concordance among the scales. Therefore, there is potential for a direct and routine translation of the current BGTS-based rule for ARIA-E-related dosing suspension in the GRADUATE trials, into a new rule based on the SSAE-3 or SSAE-5. Overall, the SSAE-3 and SSAE-5 may provide a simpler alternative to the BGTS in a clinical setting, with the SSAE-5 providing greater granularity than the SSAE-3 for characterizing radiographically moderate ARIA-E, potentially allowing for a greater spectrum of clinical intervention.

ACKNOWLEDGMENTS

The authors would like to thank all the clinical trial participants and their families. This study was funded by F. Hoffmann-La Roche Ltd. Medical writing support for the development of the manuscript

was provided by Lucy Gupta of Nucleus Global and funded by F. Hoffmann-La Roche Ltd.

CONFLICTS OF INTEREST

Gregory Klein is a full-time employee and shareholder of F. Hoffmann-La Roche Ltd. Marzia A. Scelsi is a full-time employee of Roche Products Ltd. Jerome Barakos provides both consultative services and image interpretation for Clario. Jochen B. Fiebach reports outside the submitted work personal fees from AbbVie, AC Immune, Artemida, Bioclinica/Clario, Biogen, BMS, Brainomix, Cerevast, Daiichi-Sankyo, Eisai, F. Hoffmann-La Roche AG, Eli Lilly, Guerbet, Ionis Pharmaceuticals, IQVIA, Janssen, Julius Clinical, jung diagnostics, Lysogene, Merck, Nicolab, Premier Research, and Tau Rx. Luc Bracoud is a full-time employee of Clario. Joyce Suhy is a full-time employee of Clario. Paul Delmar is a full-time employee and shareholder of F. Hoffmann-La Roche Ltd. Marco Lyons is a full-time employee and shareholder of Roche Products Ltd. Jakub Wojtowicz is a full-time employee and shareholder of F. Hoffmann-La Roche Ltd. Szofia Bullain is a full-time employee and shareholder of F. Hoffmann-La Roche Ltd. Frederik Barkhof is supported by the NIHR Biomedical Research Centre at UCLH. He is a steering committee or iDMC member for Biogen, Merck, Roche, Eisai, and Prothena; consultant for Roche, Biogen, Merck, IXICO, Janssen, and Combinostics; has research agreements with Roche, Merck, Biogen, and GE Healthcare; and is co-founder and shareholder of Queen Square Analytics LTD. Derk Purcell reports outside the submitted work personal fees from Biogen and provides both consultative services and image interpretation for Clario. Author disclosures are available in the supporting information.

REFERENCES

1. Alzheimer's Association. 2020 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2020;16(3):391-460.
2. Klein G, Delmar P, Voyle N, et al. Gantenerumab reduces amyloid- β plaques in patients with prodromal to moderate Alzheimer's disease: a PET substudy interim analysis. *Alzheimers Res Ther*. 2019;11(1):101.
3. Ostrowitzki S, Lasser RA, Dorflinger E, et al. A phase III randomized trial of gantenerumab in prodromal Alzheimer's disease. *Alzheimers Res Ther*. 2017;9(1):95.
4. Novakovic D, Feligioni M, Scaccianoce S, et al. Profile of gantenerumab and its potential in the treatment of Alzheimer's disease. *Drug Des Devel Ther*. 2013;7:1359-1364.
5. Kwan ATH, Arfaie S, Therriault J, et al. Lessons learnt from the second generation of anti-amyloid monoclonal antibodies clinical trials. *Dement Geriatr Cogn Disord*. 2020;49(4):334-348.
6. Vander Zanden CM, Chi EY. Passive immunotherapies targeting amyloid beta and tau oligomers in Alzheimer's disease. *J Pharm Sci*. 2020;109(1):68-73.
7. Biogen Inc. ADUHELM™ (aducanumab-awwa). Prescribing Information. Accessed July 20, 2022. https://www.accessdata.fda.gov/drugsatfda_docs/label/2022/761178s005lbl.pdf
8. VJdementia. Potential new therapies for Alzheimer's disease to look out for in 2022: anti-amyloid- β antibodies. Accessed July 20, 2022. <https://vjdementia.com/feature/potential-new-therapies-for-alzheimers-disease-to-look-out-for-in-2022-anti-amyloid-%CE%B2-antibodies/>
9. Sperling RA, Jack CR, Jr, Black SE, et al. Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: recommendations from the Alzheimer's Association Research Roundtable Workgroup. *Alzheimers Dement*. 2011;7(4):367-385.
10. Antolini L, DiFrancesco JC, Zedde M, et al. Spontaneous ARIA-like events in cerebral amyloid angiopathy-related inflammation: a multicenter prospective longitudinal cohort study. *Neurology*. 2021;97(18):e1809-e1822.
11. Xiong M, Jiang H, Serrano JR, et al. APOE immunotherapy reduces cerebral amyloid angiopathy and amyloid plaques while improving cerebrovascular function. *Sci Transl Med*. 2021;13(581):eabd7522.
12. Barkhof F, Daams M, Scheltens P, et al. An MRI rating scale for amyloid-related imaging abnormalities with edema or effusion. *AJNR Am J Neuroradiol*. 2013;34(8):1550-1555.
13. Bracoud L, Fiebach JB, Purcell DD, et al. P1-047: validation of a simple severity scale for assessing ARIA-E. *Alzheimers Dement*. 2017;13(7S_Part_5):P253-P254.
14. ClinicalTrials.gov. A study of gantenerumab in participants with prodromal Alzheimer's disease (Scarlet road). ClinicalTrials.gov Identifier: NCT01224106. Accessed July 20, 2022. <https://clinicaltrials.gov/ct2/show/NCT01224106>
15. ClinicalTrials.gov. A study of gantenerumab in participants with mild Alzheimer disease. ClinicalTrials.gov Identifier: NCT02051608. Accessed July 20, 2022. <https://clinicaltrials.gov/ct2/show/NCT02051608>
16. Carlson C, Siemers E, Hake A, et al. Amyloid-related imaging abnormalities from trials of solanezumab for Alzheimer's disease. *Alzheimers Dement (Amst)*. 2016;2:75-85.
17. Salloway S, Chalkias S, Barkhof F, et al. Amyloid-related imaging abnormalities in 2 phase 3 studies evaluating aducanumab in patients with early Alzheimer disease. *JAMA Neurol*. 2022;79(1):13-21.
18. Swanson CJ, Zhang Y, Dhadda S, et al. A randomized, double-blind, phase 2b proof-of-concept clinical trial in early Alzheimer's disease with lecanemab, an anti-A β protofibril antibody. *Alzheimers Res Ther*. 2021;13(1):80.
19. Mintun MA, Lo AC, Evans CD, et al. Donanemab in early Alzheimer's disease. *N Engl J Med*. 2021;384(18):1691-1704.
20. Budd Haeberlein S, Aisen PS, Barkhof F, et al. Two randomized phase 3 studies of aducanumab in early Alzheimer's disease. *J Prev Alzheimers Dis*. 2022;9(2):197-210.
21. ClinicalTrials.gov. A study to evaluate the safety, tolerability, and efficacy of long-term gantenerumab administration in participants with Alzheimer's disease (AD). ClinicalTrials.gov Identifier: NCT04374253. Accessed July 20, 2022. <https://clinicaltrials.gov/ct2/show/NCT04374253>
22. ClinicalTrials.gov. A study to evaluate the efficacy and safety of gantenerumab in participants at risk for or at the earliest stages of Alzheimer's disease (AD) (SKYLINE). ClinicalTrials.gov Identifier: NCT05256134. Accessed July 20, 2022. <https://clinicaltrials.gov/ct2/show/NCT05256134>
23. Klein G, Bracoud L, Purcell DD, et al. Calibration of a simplified ARIA-E MRI severity scale suitable for clinical practice. *Alzheimers Dement*. 2020;16(S2):e041110.
24. Barakos J, Purcell D, Suhy J, et al. Detection and management of amyloid-related imaging abnormalities in patients with Alzheimer's disease treated with anti-amyloid beta therapy. *J Prev Alzheimers Dis*. 2022;9(2):211-220.

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: Klein G, Scelsi MA, Barakos J, et al. Comparing ARIA-E severity scales and effects of treatment management thresholds. *Alzheimer's Dement*. 2022;14:e12376. <https://doi.org/10.1002/dad2.12376>