MRI Follow-up of Astrocytoma: Automated Coregistration and Color-Coding of FLAIR Sequences Improves Diagnostic Accuracy With Comparable Reading Time

Simon Lennartz, MD,[1,2,3*] David Zopfs, MD,[1] Anne Nobis,[1] Stefanie Paquet, MD,[1] Ulrike Cornelia Isabel Hoyer, MD,[1] Charlotte Zäske, MD,[1] Lukas Goertz, MD,[4] Christoph Kabbasch, MD,[1] Kai Roman Laukamp, MD,[1] Nils Große Hokamp, MD,[1] Norbert Galldiks, MD,[5,6,7] and Jan Borggrefe, MD[1]

Background: MRI follow-up is widely used for longitudinal assessment of astrocytoma, yet reading can be tedious and error-prone, in particular when changes are subtle.
Purpose/Hypothesis: To determine the effect of automated, color-coded coregistration (AC) of fluid attenuated inversion recovery (FLAIR) sequences on diagnostic accuracy, certainty, and reading time compared to conventional follow-up MRI assessment of astrocytoma patients.
Study Type: Retrospective.
Population: In all, 41 patients with neuropathologically confirmed astrocytoma.
Field Strength/Sequence: 1.0–3.0T/FLAIR
Assessment: The presence or absence of tumor progression was determined based on FLAIR sequences, contrast-enhanced T_1 sequences, and clinical data. Three radiologists assessed 47 MRI study pairs in a conventional reading (CR) and in a second reading supported by AC after 6 weeks. Readers determined the presence/absence of tumor progression and indicated diagnostic certainty on a 5-point Likert scale. Reading time was recorded by an independent assessor.
Statistical Tests: The Wilcoxon test was used to assess reading time and diagnostic certainty. Differences in diagnostic accuracy, sensitivity, and specificity were analyzed with the McNemar mid-p test.
Results: Readers attained significantly higher overall sensitivity (0.86 vs. 0.75; $P < 0.05$) and diagnostic accuracy (0.84 vs. 0.73; $P < 0.05$) for detection of progressive nonenhancing tumor burden when using AC compared to CR. There was a strong trend towards higher specificity within the AC-augmented reading, yet without statistical significance (0.83 vs. 0.71; $P = 0.08$). Sensitivity for unequivocal disease progression was similarly high in both approaches (AC: 0.94, CR: 0.92), while for marginal disease progressions, it was significantly higher in AC (AC: 0.78, CR: 0.58; $P < 0.05$). Reading time including application loading time was comparable (AC: 38.1 $\pm$ 16.8 sec, CR: 36.0 $\pm$ 18.9 s; $P = 0.25$).
Data Conclusion: Compared to CR, AC improves comparison of FLAIR signal hyperintensity at MRI follow-up of astrocytoma patients, allowing for a significantly higher diagnostic accuracy, particularly for subtle disease progression at a comparable reading time.
Evidence Level: 3
Technical Efficacy Stage: 6

J. MAGN. RESON. IMAGING 2020;52:1197–1206.

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.27136

Received Jan 17, 2020, Accepted for publication Mar 3, 2020.

*Address reprint request to: S.L, Kerpener Straße 62, 50937 Cologne, Germany Email: simon.lennartz@uk-koeln.de
Grant Support: This work was funded through the Else Kröner-Fresenius Stiftung (2016-Kolleg-19 to Simon Lennartz). The funding source hat no involvement in study design, the collection, analysis, or interpretation of data or in the writing of the report.

From the [1]Institute for Diagnostic and Interventional Radiology, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany; [2]Else Kröner Forschungskolleg Clonal Evolution in Cancer, University Hospital Cologne, Cologne, Germany; [3]Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA; [4]Center for Neurosurgery, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany; [5]Department of Neurology, Faculty of Medicine and University Hospital Cologne, University of Cologne, Köln, Germany; [6]Institute of Neuroscience and Medicine (INM-3), Research Center Juelich, Juelich, Germany; and [7]Center of Integrated Oncology (CIO), Universities of Aachen, Bonn, Cologne and Duesseldorf, Germany

ASTROCYTOMA is the most common histologic subtype of glioma[1] and can be subdivided into circumscribed or diffuse astrocytic tumors on the one hand and low- or high-grade astrocytoma on the other. According to the 2007 WHO classification, low-grade astrocytomas predominantly arise in younger adults and show a rather slow growth rate,[2] while high-grade astrocytomas primarily occur in middle-aged individuals and usually grow more rapidly.[3] Subsequently, it could be demonstrated that predominantly the isocitrate dehydrogenase (IDH) mutational status is the most important determinant of prognosis.[4,5]

Magnetic resonance imaging (MRI) is the modality of choice for imaging of brain tumors[6] and for follow-up of patients in the pre- and posttherapeutic stage.[7] For the latter, contrast-enhanced T_1-weighted images are key, as higher tumor grades are strongly associated with gadolinium contrast enhancement.[8] However, contrast enhancement may also be absent in high-grade glioma, ie, following antiangiogenic treatment, and is uncommon in low-grade tumors.[8,9] Therefore, fluid attenuated inversion recovery (FLAIR) sequences are integral for MRI follow-up of astrocytoma: according to the Response Assessment in Neuro-Oncology (RANO) criteria, a significant increase in FLAIR hyperintensity warrants determination of disease progression for high-grade glioma. For low-grade glioma, FLAIR is the key sequence to determine disease state, as they rarely show contrast enhancement.[10,11] Assessing repetitive MRI examinations, however, is time-consuming, and comparability may be hampered by discrepant acquisition protocols or the use of heterogeneous MRI scanners. Thus, techniques for subtraction and automated coregistration of such repetitive follow-ups have been developed and reported to be useful, particularly for follow-up of multiple sclerosis (MS) with high lesion burden.[12,13] It was reported that along with the use of these tools, diagnostic accuracy could be increased, while required reading was reduced.[14]

However, for the purpose of follow-up imaging of heteromorphic brain tumors as astrocytoma, for which longitudinal evaluation of disease is paramount, these tools have not been systematically investigated yet. Our hypothesis was that an automated coregistration approach could yield comparable advantages for astrocytoma follow-up, as previously shown for MS patients.[12,13] Hence, the study purpose was to compare diagnostic accuracy, required reading time, and diagnostic certainty between longitudinal assessment of FLAIR sequences with and without automated coregistration in patients with low- or high-grade astrocytoma.

Materials and Methods

Patient Acquisition

The Institutional Review Board approved this study and waived informed consent. Eligible patients were retrospectively identified

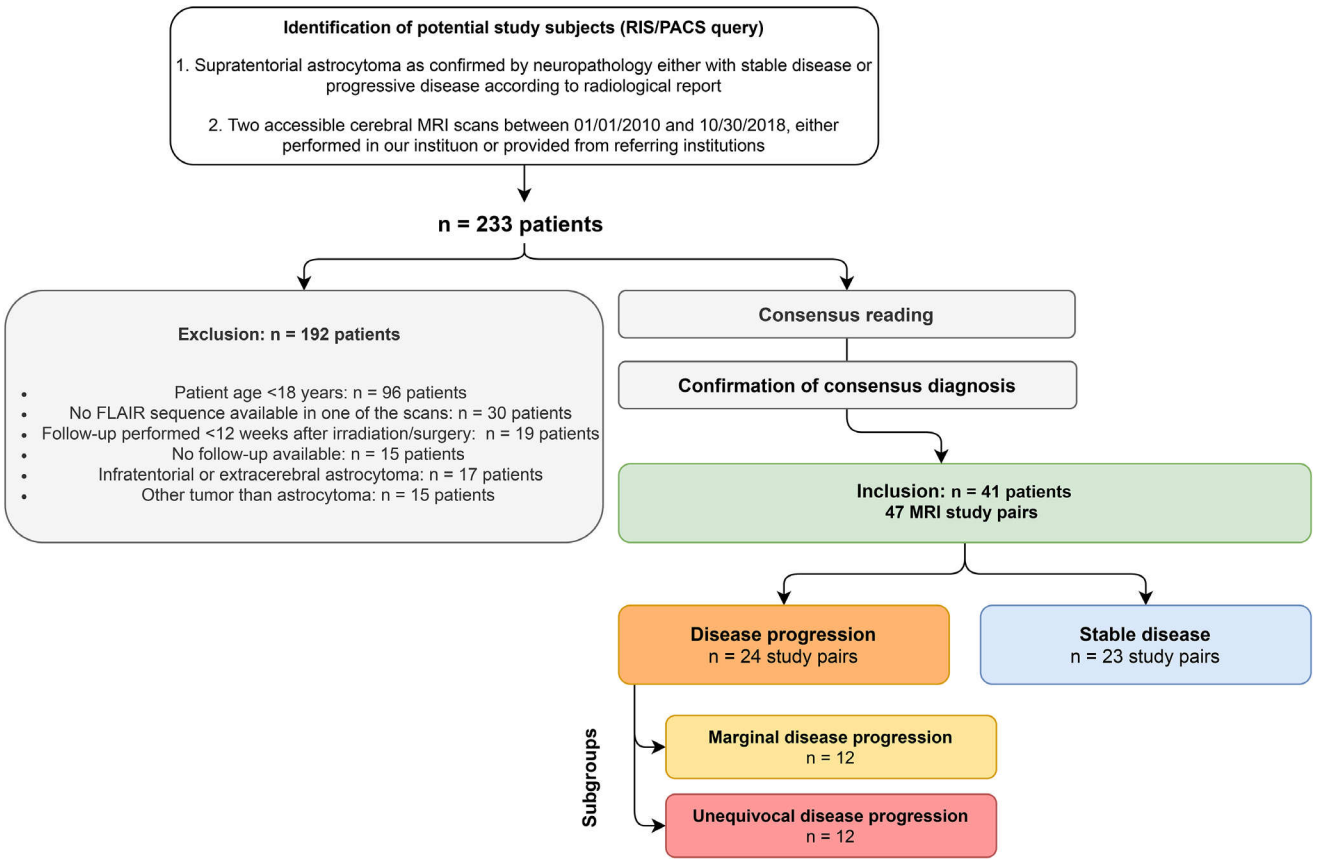

FIGURE 1: Workflow for inclusion and exclusion of study subjects.

TABLE 1. Overview of Scanner Parameters Used for Scans Performed in Our Institution and at Referring Institutions Showing Ranges of Field of View, Matrix, Slice Thickness, Repetition Time (TR), Echo Time (TE), and Inversion Time (TI)

	Philips Achieva 3.0T	Philips Ingenia 3.0T	Philips Achieva 1.5T	Philips Ingenia 1.5T
Field of view (mm)	512 × 512–560 × 560	512 × 512–1024 × 1024	256 × 256	256 × 256–448 × 448
Matrix	256 × 201–280 × 211	224 × 223–280 × 211	256 × 172	256 × 177–256 × 200
Slice thickness (mm)	2–5	1–5	4	2–5
TR (msec)	12000	12000	6000	6000
TE (msec)	140	140	120	140
TI (msec)	2850	2850	2000	2000

	Philips Intera 1.5T	Philips Intera 1.5T	Philips Panorama 1.0T	Scanners of referring institutions (Toshiba, Philips [0.5–1.5T])
Field of view (mm)	256 × 256	256 × 256	256 × 256–288 × 288	256 × 256–640 × 640
Matrix	256 × 203	256 × 203	256 × 154–280 × 172	256 × 173–320 × 193
Slice thickness (mm)	6	6	6	5–6
TR (msec)	6000	6000	6000	5000–7200
TE (msec)	100	100	120	100–120
TI (msec)	2000	2000	2000	1900–2200

with a combined database query to the radiological information and picture archiving and communication systems (RIS/PACS).

Included were patients diagnosed with supratentorial astrocytoma as confirmed by an unequivocal neuropathologic report as low- or high-grade astrocytoma either with stable disease or progressive disease according to consensus reading and ground truth correlation. Moreover, these patients needed to have undergone two cerebral MRI scans between 01/01/2010 and 10/30/2018 accessible through PACS, either performed in our institution or provided from referring institutions. Patients who were under 18 years, who did not have a FLAIR sequence available in one of the scans, or whose follow-up MRI had been performed less than 12 weeks after irradiation/surgery were excluded, the latter to minimize the risk of including cases with pseudoprogression.[10] A detailed patient acquisition workflow is shown in Fig. 1.

Image Acquisition

The provided MRI data were obtained from different vendors, scanner types/generations, and included different field strengths (1.0T/1.5T/3.0T). Table 1 provides an overview of the detailed scan parameters. In total, 94 datasets were included, of which 87 were acquired in our institution and 7 in referring institutions.

Ground Truth Annotation

Ground truth (ie, disease progression with increasing FLAIR signal hyperintensity or stable disease with constant FLAIR signal hyperintensity) was determined in consensus by two radiologists, of which one was a senior, board-certified neuroradiologist (S.L., 3 years of experience, J.B., 11 years of experience). The annotation was primarily carried out in PACS and substantiated by information from the automated coregistration (AC) software. To warrant consistency between this imaging-based diagnosis and actual disease status, all clinical and additional imaging data were accessible. If a case was determined as progression, this progression was to be confirmed based on either a further, disease-related increase in FLAIR hyperintensity in subsequent follow-up MRI or neuropathologic confirmation of disease progression after subsequent biopsy/resection.

Subgroup Analysis

To facilitate subgroup analysis of marginal and unequivocal progression of nonenhancing tumor burden, FLAIR signal hyperintensity was measured bidirectionally and the product of these bidirectional measurements was calculated as suggested for evaluation according to the RANO criteria. Marginal and unequivocal disease progressions were considered as the lower and upper quartiles of the proportional change of the referring products between follow-ups, respectively. Additionally, disease state according to the RANO criteria was recorded: for low-grade astrocytomas, the diagnosis was based on the calculated change of the product of bidirectional measurements of FLAIR hyperintensity between the two examinations,[15] whereas for high-grade astrocytomas disease state was determined visually.[10] Figure 2 shows examples for stable and progressive disease as determined by the RANO criteria.

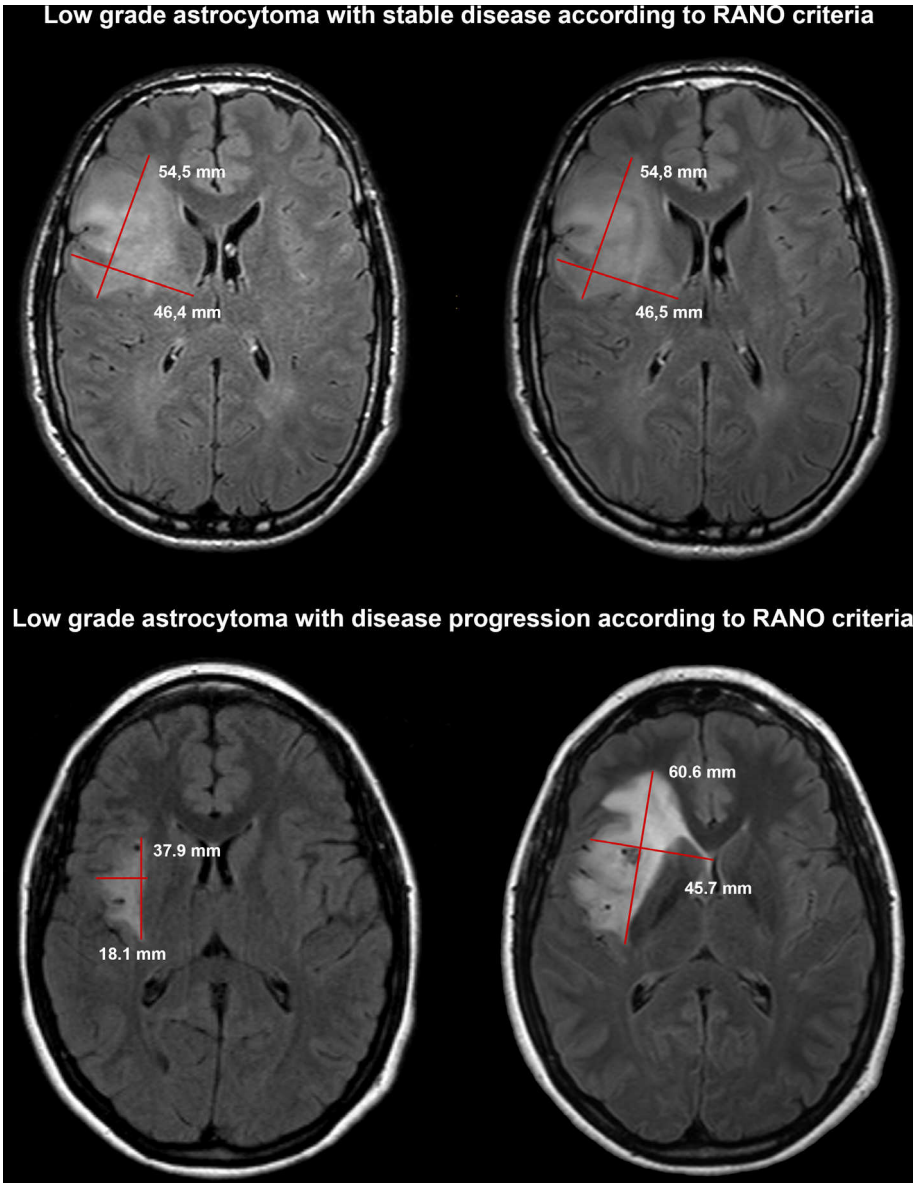

FIGURE 2: Tumor measurements to determine the presence or absence of disease progression according to the RANO criteria.

Subjective Assessment

Three radiologists with different levels of expertise in neuroimaging (3, 6, and 13 years of experience) who were blinded with regard to patient characteristics, examination date, tumor grade, and any other clinical data independently assessed the included 47 MRI study pairs in a two-step procedure: First, the images were evaluated using the conventional reading (CR) approach. Second, after a time interval of 6 weeks that was set to reduce recall bias, the images were assessed again in the AC session using a different, randomized order. In both sessions the subjective assessors determined disease status (progressive disease / stable disease) and indicated diagnostic certainty for each decision. For both reading sessions, a timekeeper recorded the time in seconds needed for the individual cases (from the point at which the images were completely loaded until both decisions had been reported). Additionally, loading times encountered for the AC software and the PACS were determined.

Conventional and Automated Coregistration Reading

For the CR approach, the local PACS software was used, which is the standard tool for follow-up MRI assessment in clinical routine at our institution (Impax EE R20, XVII SU1, Agfa Healthcare, Mortsel, Belgium). The axial FLAIR sequences for the different timepoints were arranged side-by-side in advance. The raters were free to manually link the images at corresponding slices and to adjust window settings, as in clinical practice.

AC reading was carried out with a CE-certified and FDA-approved AC tool (LoBI, Philips Healthcare, Best, Netherlands) which is integrated in the vendor's image viewer (Intellispace Portal 11, Philips Healthcare). The AC tool comprises a multistep image processing pipeline to correct for differences between the two datasets that are supposed to be compared: First, it executes a bias field correction, after which the bias field-corrected images undergo a rigid coregistration. Last, an intensity scaling is carried out to

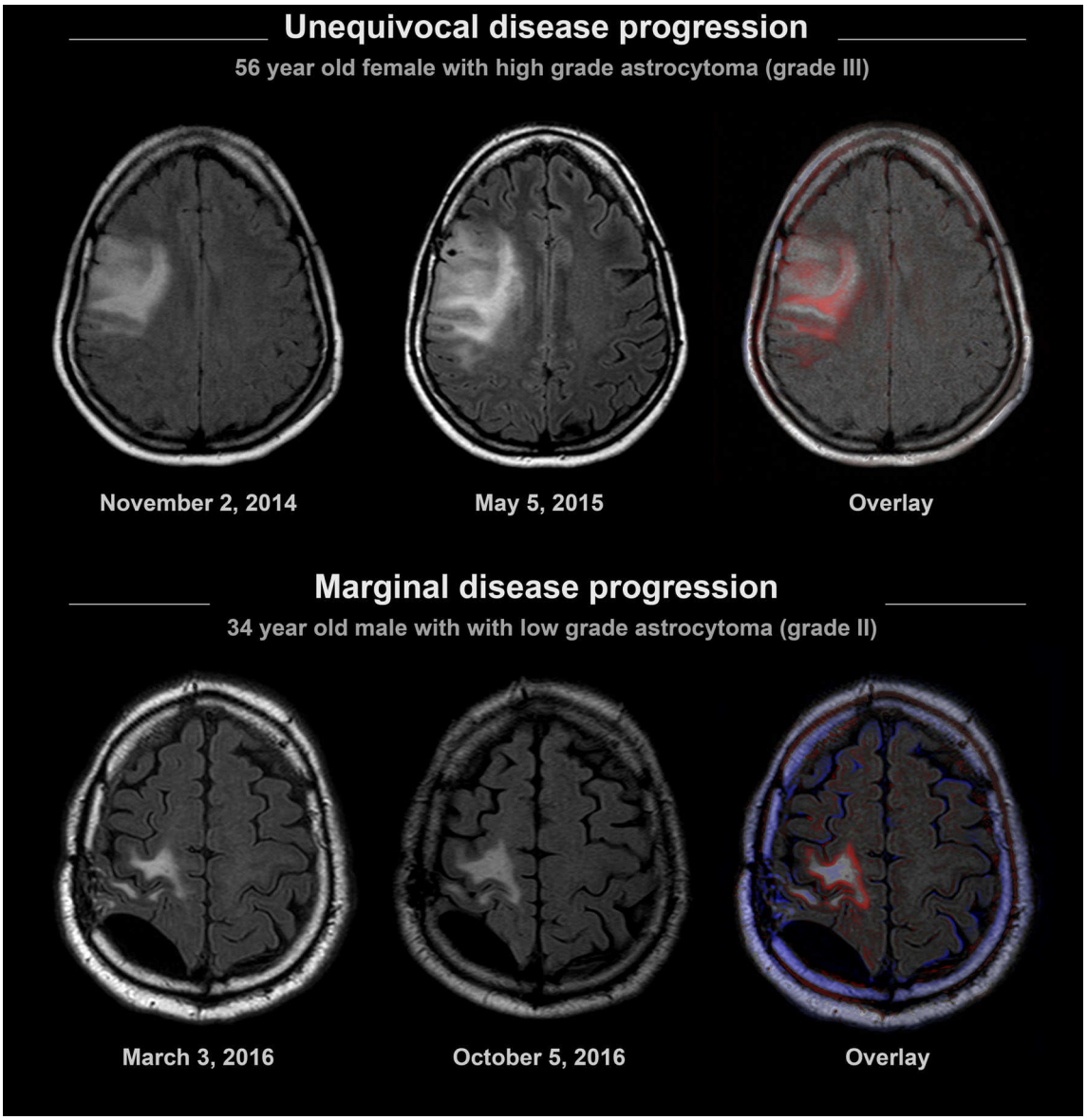

FIGURE 3: Examples of marginal and unequivocal progression of nonenhancing tumor burden. Both cases show prior scans (left image), follow-ups (center image), and automatically coregistered, color-coded overlays of both (right image) indicating progression of FLAIR signal hyperintensity in red.

correct for possible differences in signal intensity that would impede depiction of actual, disease-related changes in signal hyper- or hypointensity. These changes are then color-coded so that a progression of FLAIR hyperintensity is indicated as red, while a regression is marked blue on the referring overlay map (Fig. 3).

After selection of the follow-up study pairs, the color-coded overlay map is presented to the radiologist in a side-by-side setup along with the two original gray-scaled images so that each possible change in signal that is depicted on the overlay image can be correlated with the original images on the same anatomical level.

Statistical Analysis

Statistical evaluations were performed with JMP software (v. 14, SAS Institute, Cary, NC) and Microsoft Excel (Microsoft, Redmond, WA). The Wilcoxon test was used to determine statistical differences between reading time and diagnostic certainty. The McNemar mid-p test was used to account for differences in diagnostic accuracy, sensitivity, and specificity.[16] Statistical significance was defined as $P < 0.05$. Agreement among the three subjective readers was evaluated using the intraclass correlation coefficient (ICC) and interpreted as suggested earlier[17]: poor: <0.40; fair: 0.41–0.59; good: 0.6–0.74; excellent: 0.75–1.00. Continuous variables are indicated as mean ± standard deviation, while ordinal variables are indicated as median and interquantile range.

Results

Patients

In total, 41 patients were included, of which 21 were men and 20 were women. Thirty-five patients contributed one study pair each, whereas six contributed two study pairs each, resulting in 47 study pairs that were finally included. The

TABLE 2. Patient Characteristics Showing Cases With Disease Progression According to Ground Truth Annotation and Response Assessment in Neuro-Oncology (RANO) Criteria

Patient characteristics	
WHO tumor grade as confirmed by neuropathology	(n = 41 patients)
Low-grade astrocytoma	16 (39%)
High-grade astrocytoma	25 (61%)
Radiological diagnosis as determined in consensus reading and confirmed by clinical information/follow-up	(n = 47 study pairs)
Disease progression with increasing FLAIR signal hyperintensity	24 (51%)
Unequivocal disease progressions (upper two quartiles of FLAIR signal hyperintensity change in %)	12 (50%)
Marginal disease progressions (lower two quartiles of FLAIR signal hyperintensity change in %)	12 (50%)
Stable disease with constant FLAIR signal hyperintensity	23 (49%)
Evaluation according to RANO criteria	(n = 47 study pairs)
Disease progression	20 (43%)
Stable disease	27 (57%)

mean age was 44.5 ± 15.5 years. Of the 41 patients included, 25 were diagnosed with high-grade astrocytoma and 16 with low-grade astrocytoma. In 24 of the 47 study pairs, a progressive FLAIR signal hyperintensity were diagnosed in CR, while in 23 follow-ups the FLAIR signal hyperintensity was determined stable. In four cases of low-grade astrocytoma that were determined as "stable" according to the RANO criteria, a progressive FLAIR signal hyperintensity were found in CR that was associated with a true tumor progression following ground truth correlation. Table 2 gives an overview of underlying diseases.

Subjective Assessment

Overall sensitivity and diagnostic accuracy for progression of FLAIR signal hyperintensity were significantly higher in the reading supported by AC than in the CR approach (sensitivity/accuracy: 0.86/0.84 vs. 0.75/0.73; $P < 0.05$; Fig. 4). Specificity was markedly higher for AC (0.83) than for CR (0.71) as well, yet without statistical significance ($P = 0.12$). Sensitivity for marginal progressive diseases significantly increased from 0.58 in CR to 0.78 in the AC reading ($P < 0.05$). In contrast, sensitivity for unequivocal progressive diseases was on a similarly high level in both reading approaches (0.92 vs. 0.94; see also Fig. 5). Diagnostic accuracy found in the AC reading for follow-up pairs with a slice thickness difference greater than 2 mm was 0.79. Overall interobserver agreement regarding detection of disease progression was poor (ICC = 0.32) in the conventional reading and markedly improved in the AC approach, where it was good (ICC = 0.72).

Diagnostic certainty for all cases was higher in AC reading (5 (4–5)) than in CR (4 (3–5); $P < 0.05$). For the subgroup of unequivocal progressive cases, overall certainty was very high, so that no significant difference between both reading approaches could be observed (both 5 (5–5), $P = 1.0$). In the subgroup of marginally progressive cases, there was a tendency towards improved certainty scores in the AC reading (AC: 4 (3.25–5) vs. CR: 4 (3–5), see Fig. 6) which, however, was not statistically significant ($P = 0.7$). Diagnostic accuracy at the AC reading found in study pairs comprising scans with a slice thickness difference greater than 2 mm (n = 13) as well as different scanners/field strengths (n = 11) was on a comparable level as overall diagnostic accuracy found for all cases (0.79 and 0.85, respectively, vs. 0.84 in all study pairs; see also Table 3).

Averaged over all cases, reading time for the conventional approach was 33.5 ± 18.9 sec, which was significantly longer than an average of 21.6 ± 16.8 sec that was required in the AC ($P < 0.001$; Fig. 7). For cases with marginal disease progression, reading time required for conventional assessment was highest with 38.5 ± 16.7 sec, which was reduced to 26.2 ± 18.8 sec in the reading supported by the AC software ($P < 0.05$). However, these time savings were neutralized by the longer average loading time of 16.5 sec encountered in the AC software (vs. 2.5 sec in the PACS). Hence, overall reading time implying application loading time was comparable between the two approaches for all cases (CR: 36.0 ± 18.9, AC: 38.1 ± 16.8 sec, $P = 0.25$) as well as marginally progressive cases (CR: 41.0 ± 16.7, AC: 42.7 ± 18.8 sec, $P = 0.83$). For unequivocal disease

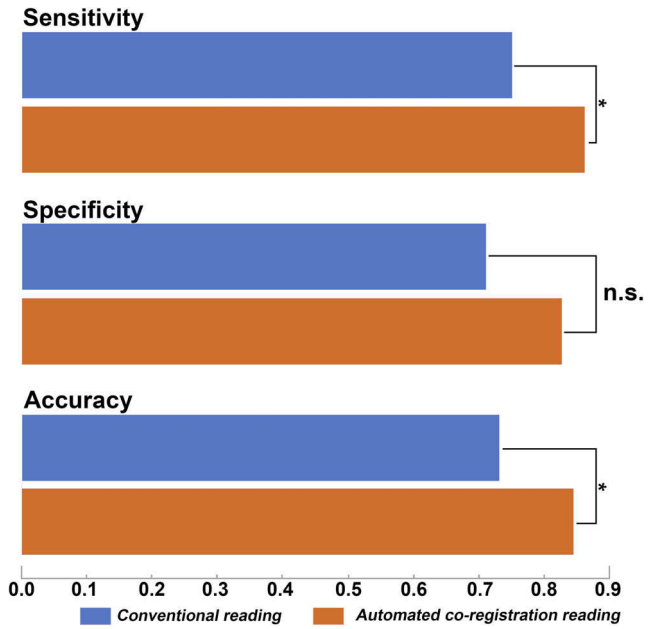

FIGURE 4: Overall sensitivity, specificity, and diagnostic accuracy for progression of FLAIR hyperintensity attained by the three readers in the assessment of supratentorial astrocytoma were significantly higher in the automated coregistration reading (orange bars) than in the conventional reading approach (blue bars). Asterisks indicate statistical significance.

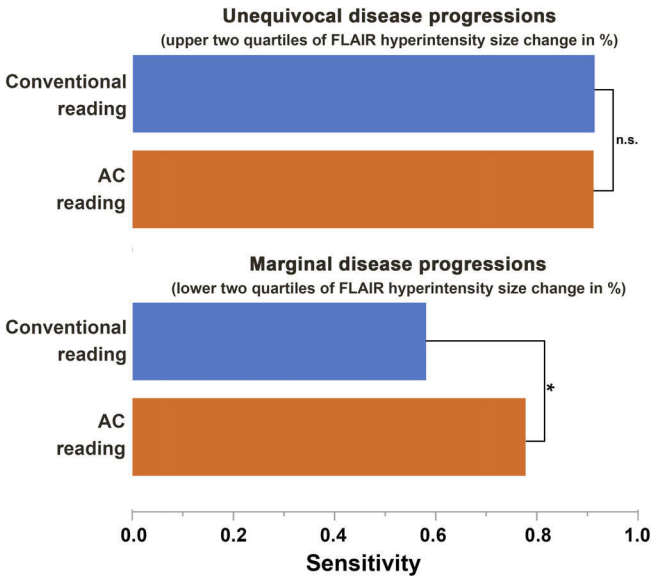

FIGURE 5: Subgroup analysis of marginal and unequivocal disease progressions as determined by size change of FLAIR hyperintensity. It is evident that automated coregistration significantly improved sensitivity for marginal disease progressions, while for unequivocal disease progressions, sensitivity was comparably high in both reading approaches. Asterisks indicate statistical significance.

progressions, time saving attained with the AC software was relatively low (CR: 16.4 ± 8.7 sec vs. AC: 11.9 ± 10.5 sec, $P < 0.05$). Hence, when implying loading times, overall assessment time for these unequivocal cases was significantly higher with the AC software (28.4 ± 10.5 sec) than with the PACS (18.9 ± 8.7 sec, $P < 0.05$).

Discussion

In the era of precision oncology, it is an ongoing challenge for diagnostic radiologists to provide a high level of accuracy with the limited time available due to steadily increasing exam numbers.[18] Among others, MRI follow-up of tumors of the central nervous system has become increasingly important to provide reliable endpoints for both existing and future therapies.[19,20] Regarding longitudinal assessment of patients with astrocytoma, a frequent brain tumor of high clinical significance,[1] assessment of FLAIR signal hyperintensity is important for follow-up of low- and high-grade tumors, since progression of nonenhancing tumor tissue is known to occur earlier than clinical progression.[10,21] Changes in nonenhancing tumor burden as depicted in FLAIR or T_2-weighted images can be subtle, making precise assessment of referring follow-up MRI exams difficult and time-consuming. Hence, in this study we aimed to investigate whether automated, color-coded coregistration (AC) of FLAIR sequences could improve diagnostic accuracy in the assessment of astrocytoma and to determine its influence on reading time and diagnostic certainty.

AC of astrocytoma follow-up study pairs allowed readers to improve their sensitivity and diagnostic accuracy for disease progression as indicated by increasing FLAIR signal hyperintensity compared to the conventional reading. Notably, interreader agreement increased from a rather poor level for the conventional reading approach to a good one in the AC reading, which supports the improved diagnostic performance and which would be beneficial when being applied clinically. This study comprised examinations from different MRI scanners and with different slice thicknesses, which can result in partial volume effects that may hamper image postprocessing methods as investigated in this study. However, we found that both for study pairs with a slice thickness difference greater than 2 mm (n = 13) and with different field strengths (1.5T/3T, n = 11), the diagnostic accuracy in the AC reading was on a similar level as when averaged over all examinations (0.79 and 0.85, respectively, vs. 0.84 for all study pairs). This indicates that coregistration and resampling still work acceptably in these cases, which would facilitate clinical application at which comparing scans from different scanners is a commonly encountered scenario. Yet the influence of differences in scan parameters on coregistration quality should be investigated in subsequent studies with a larger cohort to retrieve more robust data in this regard.

Automated coregistration tools for follow-up MRI have been investigated before, yet the main focus of previous studies was set on the longitudinal assessment of multiple sclerosis (MS) lesions.[12–14,22] In contrast, studies on the application of such tools in the field of neuro-oncology are lacking. The

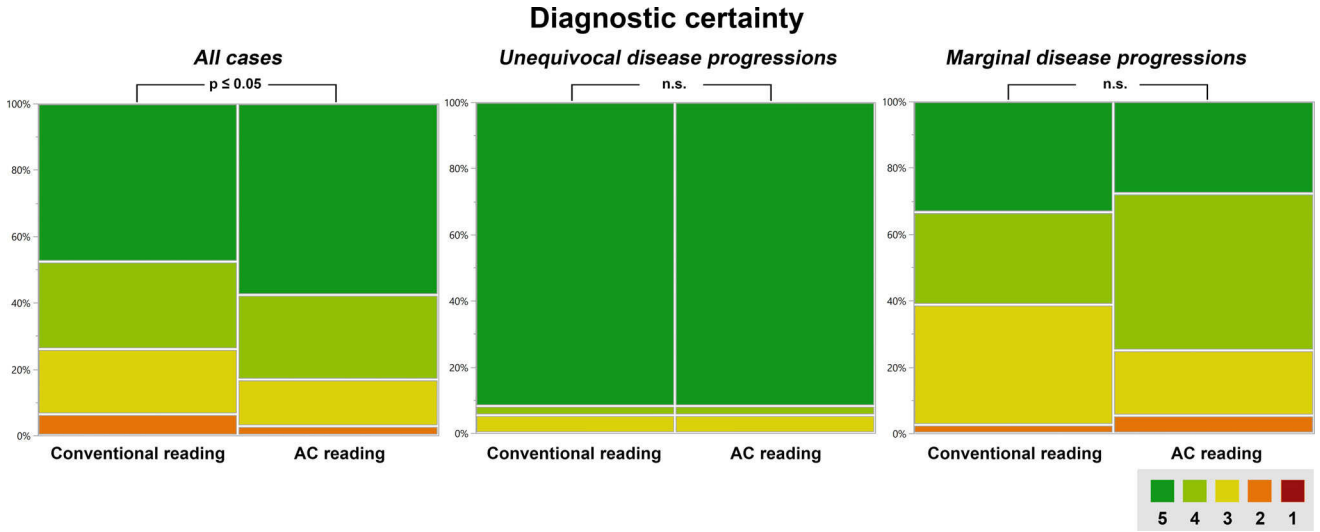

FIGURE 6: Diagnostic certainty as indicated by the readers on a 5-point Likert scale. Pertaining to all cases, certainty was significantly higher in the automated coregistration reading than in the conventional reading approach, whereas for unequivocal and marginal disease progressions, no significant difference could be found, although there was a tendency towards higher certainty scores in the AC reading for the latter.

TABLE 3. Characteristics of Study Pairs Including Proportion of Scan Pairs With Consistent Scan Parameters, Distribution of Stable Disease/Disease Progressions, as Well as Diagnostic Performance for Each Subgroup of Scanner Combinations

Scanner combination	Number of study pairs	Consistent scan parameters[a]	Stable disease	Marginal progression	Unequivocal progression	Sensitivity (CR/AC)	Specificity (CR/AC)	Accuracy (CR/AC)
3.0T/3.0T	16	14/16	7	4	5	0.67/0.85	0.76/0.71	0.71/0.79
1.5T/1.5T	14	14/14	8	3	3	0.83/0.94	0.71/0.96	0.76/0.95
1.5T/3.0T	11	0/11	5	4	2	0.78/0.94	0.60/0.73	0.70/0.85
Other	6	2/6	3	1	2	0.78/0.56	0.78/0.89	0.78/0.72
All	47	30/47	23	12	12	0.75/0.86	0.71/0.83	0.73/0.84

[a]Repetition time, echo time, inversion time, slice thickness.

initial data we present on the longitudinal assessment of astrocytoma using AC software indicate comparable advantages as revealed by the referring studies on MS follow-up, in which an increased accuracy for disease progression was reported. However, in contrast to a previous study on MS follow-up by Galletto Pregliasco et al, in which the reading time was reported to be significantly reduced, we found a comparable reading time between the AC and conventional reading approach.[12] While this might be due to software-related differences, Zopfs et al likewise reported a significant reduction of reading time in MS follow-up applying the same AC software as was used in our study.[14] Hence, it seems warranted to conclude that the potential time savings provided by AC are greater for numerous lesions such as is prevalent in MS follow-up as it is for longitudinal assessment of single, heteromorphic tumor structures, as investigated in our study.

According to the RANO criteria, for low-grade glioma a proportional increase in FLAIR signal hyperintensity has to be ≥25% to diagnose disease progression, while any significant increase in FLAIR signal hyperintensity warrants diagnosis of progression in high-grade glioma. These criteria are primarily meant to be applied in clinical trials and to prevent unwarranted exclusion or therapy discontinuations of patients with moderate FLAIR progressions, which might be due to prior radiation or antiangiogenic therapy.[9,23,24] However, in clinical routine, subtle changes in FLAIR signal hyperintensity may indicate disease progression and lead to changes of therapy or surveillance. Accordingly, in our study four cases of low-grade astrocytoma that were determined as "stable" according to the RANO criteria showed an increase in FLAIR signal hyperintensity that was associated with a true tumor progression. We found that sensitivity for these

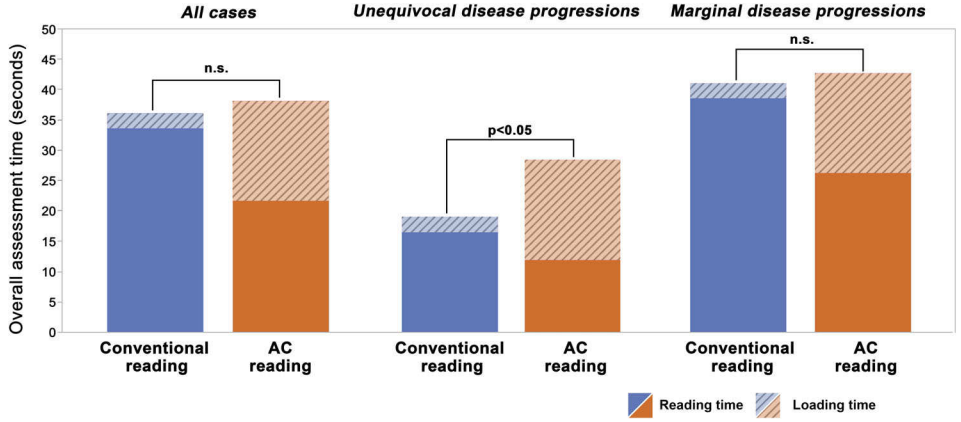

FIGURE 7: Time required to examine cases in the conventional reading (CR) as compared to the image reading with automated coregistration (AC) including the time to load (blue striped) and coregister (orange striped) MRI study pairs. Higher loading time encountered with the AC tool neutralized significant savings in reading time observed for all cases and marginal disease progressions, leading to comparable overall assessment times. For unequivocal disease progressions, overall assessment time was significantly higher when using the AC software, as the savings in reading time provided by the AC software were minimal.

marginal disease progressions was significantly higher when using the AC software. Opposingly, in the case of unequivocal disease progression, there was, as expected, no difference between the two reading approaches, yet a significantly higher overall assessment time for the AC reading. This shows that the particular strength of such AC tools lies in the depiction of rather subtle cases, while unequivocal cases can be detected in the PACS with similar accuracy in a shorter overall assessment time (Fig. 7). However, averaged over all cases, there was no significant difference in overall reading time between both approaches.

Limitations

Apart from its retrospective character, this study has limitations that need to be addressed. First, the sample size included in this study was relatively small, as rigid exclusion criteria were applied to warrant inclusion of true tumor progressions exclusively. While we did not find differences in scan parameters to have a negative impact on AC quality in our cohort, larger-scale studies in this regard should be encouraged. Second, despite the latency period of 6 weeks that was kept between the CR and AC reading, a certain recognition bias cannot entirely be excluded. Third, while FLAIR evaluation plays an important role in assessing brain tumors, our study did not investigate multiparametric image assessment in clinical routine; hence, the advantages shown for the follow-up assessment of FLAIR hyperintensity may not be applicable to other sequences of interest, such as contrast-enhanced image sequences. Fourth, our study only comprised adult patients, which limits the generalizability of our results to infants and children. Last, the technique we investigated does not address the issue of differentiation between true tumor progression and therapy-related changes in FLAIR signal associated with pseudoprogressions, for which more advanced image interpretation techniques such as radiomics may be advantageous.[25] However, clinical integration and validation of such methods can be considered more difficult than for the technique we investigated, which clearly improved visualization of subtle changes in FLAIR hyperintensity and may therefore be beneficial for detecting such changes in clinical routine.

Conclusion

This study showed that automated, color-coded coregistration facilitated longitudinal assessment of FLAIR sequences in astrocytoma patients, leading to a significantly higher sensitivity and diagnostic accuracy for progression of nonenhancing tumor with comparable reading time. Hence, especially in the light of markedly improved detection of marginal disease progressions, clinical implementation of such tools seems advantageous for longitudinal assessment of nonenhancing tumor burden of astrocytoma patients.

Acknowledgments

We thank Frank Thiele for providing helpful information on technical questions.

References

1. Ostrom QT, Bauchet L, Davis FG, et al. The epidemiology of glioma in adults. A "state of the science" review. Neuro Oncol 2014;16:896-913.
2. Forst DA, Nahed BV, Loeffler JS, Batchelor TT. Low-grade gliomas. Oncologist 2014;19:403-413.
3. Hathout L, Pope WB, Lai A, Nghiemphu PL, Cloughesy TF, Ellingson BM. Radial expansion rates and tumor growth kinetics predict malignant transformation in contrast-enhancing low-grade diffuse astrocytoma. CNS Oncol 2015;4:247-256.

4. Reuss DE, Sahm F, Schrimpf D, et al. ATRX and IDH1-R132H immunohistochemistry with subsequent copy number analysis and IDH sequencing as a basis for an "integrated" diagnostic approach for adult astrocytoma, oligodendroglioma and glioblastoma. Acta Neuropathol 2015;129:133-146.
5. Louis DN, Perry A, Reifenberger G, et al. The World Health Organization classification of tumors of the central nervous system. A summary. Acta Neuropathol 2016;2016(131):803-820.
6. Fouke SJ, Benzinger T, Gibson D, Ryken TC, Kalkanis SN, Olson JJ. The role of imaging in the management of adults with diffuse low grade glioma. A systematic review and evidence-based clinical practice guideline. J Neurooncol 2015;125:457-479.
7. Villanueva-Meyer JE, Mabray MC, Cha S. Current clinical brain tumor imaging. Neurosurgery 2017;81:397-415.
8. Lote K, Egeland T, Hager B, Skullerud K, Hirschberg H. Prognostic significance of CT contrast enhancement within histological subgroups of intracranial glioma. J Neurooncol 1998;40:161-170.
9. Nowosielski M, Wiestler B, Goebel G, et al. Progression types after antiangiogenic therapy are related to outcome in recurrent glioblastoma. Neurology 2014;82:1684-1692.
10. Wen PY, Chang SM, van den Bent MJ, Vogelbaum MA, Macdonald DR, Lee EQ. Response assessment in neuro-oncology clinical trials. J Clin Oncol 2017;35:2439-2449.
11. Chukwueke UN, Wen PY. Use of the response assessment in neuro-oncology (RANO) criteria in clinical trials and clinical practice. CNS Oncol 2019;8:CNS28.
12. Galletto Pregliasco A, Collin A, Guéguen A, et al. Improved detection of new MS lesions during follow-up using an automated MR coregistration-fusion method. AJNR Am J Neuroradiol 2018;39:1226-1232.
13. Patel N, Horsfield MA, Banahan C, et al. Detection of focal longitudinal changes in the brain by subtraction of MR images. AJNR. Am J Neuroradiol 2017;38:923-927.
14. Zopfs D, Laukamp KR, Paquet S, et al. Follow-up MRI in multiple sclerosis patients. Automated co-registration and lesion color-coding improves diagnostic accuracy and reduces reading time. Eur Radiol 2019;29(12):7047-7054. [epub ahead of print].
15. van West SE, de Bruin HG, van de Langerijt B, Swaak-Kragten AT, van den Bent MJ, Taal W. Incidence of pseudoprogression in low-grade gliomas treated with radiotherapy. Neuro Oncol 2017;19:719-725.
16. Fagerland MW, Lydersen S, Laake P. The McNemar test for binary matched-pairs data. Mid- p and asymptotic are better than exact conditional. BMC Med Res Methodol 2013;13:1-8.
17. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6:284-290.
18. Loggers ET, Fishman PA, Peterson D, et al. Advanced imaging among health maintenance organization enrollees with cancer. J Oncol Pract 2014;10:231-238.
19. Wen PY, Cloughesy TF, Ellingson BM, et al. Report of the Jumpstarting Brain Tumor Drug Development Coalition and FDA clinical trials neuroimaging endpoint workshop (January 30, 2014, Bethesda MD). Neuro Oncol 2014;16(Suppl 7):36-47.
20. van den Bent MJ, Vogelbaum MA, Wen PY, Macdonald DR, Chang SM. End point assessment in gliomas. Novel treatments limit usefulness of classical Macdonald's criteria. J Clin Oncol 2009;27:2905-2908.
21. Kazda T, Hardie JG, Pafundi DH, Kaufmann TJ, Brinkmann DH, Laack NN. Evaluation of RANO response criteria compared to clinician evaluation in WHO grade III anaplastic astrocytoma. Implications for clinical trial reporting and patterns of failure. J Neurooncol 2015;122:197-203.
22. van Heerden J, Rawlinson D, Zhang AM, et al. Improving multiple sclerosis plaque detection using a Semiautomated assistive approach. AJNR Am J Neuroradiol 2015;36:1465-1471.
23. Thust SC, van den Bent MJ, Smits M. Pseudoprogression of brain tumors. J Magn Reson Imaging 2018;9:571-589. [epub ahead of print].
24. Brandsma D, van den Bent MJ. Pseudoprogression and pseudoresponse in the treatment of gliomas. Curr Opin Neurol 2009;22:633-638.
25. Chen X, Wei X, Zhang Z, Yang R, Zhu Y, Jiang X. Differentiation of true-progression from pseudoprogression in glioblastoma treated with radiation therapy and concomitant temozolomide by GLCM texture analysis of conventional MRI. Clin Imaging 2015;39:775-780.