



Interrater reliability of Brain Tumor Reporting and Data System (BT-RADS) in the follow up of adult primary brain tumors: a single institution experience in Italy

Marco Parillo^{1,2^}, Carlo Augusto Mallio^{1,2}, Matteo Pileri^{1,2}, Diab Dirawe^{1,2}, Andrea Romano³, Alessandro Bozzao³, Brent Weinberg⁴, Carlo Cosimo Quattrocchi^{1,2}

¹Unit of Diagnostic Imaging and Interventional Radiology, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Roma, Italy; ²Fondazione Policlinico Universitario Campus Bio-Medico, Roma, Italy; ³Neuroradiology Unit, NESMOS Department (Neuroscience, Mental Health, Sense Organs), Sant'Andrea Hospital, Medical and Psychology School, La Sapienza University, Rome, Italy; ⁴Neuroradiology Division, Department of Radiology and Imaging Sciences, Emory University School of Medicine, Atlanta, GA, USA

Contributions: (I) Conception and design: M Parillo, CC Quattrocchi; (II) Administrative support: CC Quattrocchi; (III) Provision of study materials or patients: M Parillo, CC Quattrocchi, CA Mallio; (IV) Collection and assembly of data: M Parillo, CA Mallio, M Pileri, D Dirawe; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Carlo Cosimo Quattrocchi, MD, PhD. Centre for Medical Sciences-CISMed, University of Trento, Via S. Maria Maddalena 1, 38122 Trento, Italy. Email: carlo.quattrocchi@unitn.it.

Background: In 2018, a new system was proposed for classifying and reporting post-treatment adult brain tumor on magnetic resonance imaging, named as Brain Tumor Reporting and Data System (BT-RADS), that needs a validation by means of agreement studies.

Methods: A retrospective study was designed with the aim of identifying contrast-enhanced magnetic resonance imaging (MRI) of adult patients on follow-up for primary brain tumor at Fondazione Policlinico Campus Bio-Medico. Four radiologists (2 radiology residents, 1 general radiologist, 1 neuroradiologist) read and scored each study using the BT-RADS scoring tool, blinded to the MRI original report. Interobserver agreement and Fleiss' k were calculated to assess the level of diagnostic agreement. It was assessed how many times the assignment of different scoring of BT-RADS would have led to a different patient management.

Results: The total number of patients included in the study was 23 with 147 MRIs and a total of 588 BT-RADS scores retrospectively evaluated. The two most frequent tumor types were astrocytoma grade 4 (62%) and oligodendroglioma grade 3 (21%). The overall agreement rate for all 4 radiologists was 82% with a Fleiss' k of 0.70. The overall agreement rate between general radiologist and neuroradiologist was 91% with a Fleiss' k of 0.86. The overall agreement rate between 2 radiology residents and neuroradiologist was 80% with a Fleiss' k of 0.66. Astrocytoma grade 3 (k : 0.51) and oligodendroglioma grade 2 (k : 0.32) showed a poor agreement while higher values of agreement were found for astrocytoma grade 4 (k : 0.70), astrocytoma grade 2 (k : 0.78) and oligodendroglioma grade 3 (k : 0.78). All the radiologists agreed on BT-RADS assignment in 70% patients, three radiologists agreed in 17% and two radiologists agree in 13%. In no cases there was a complete disagreement among the readers. In 18% of cases the discrepancy in the estimated BT-RADS would have led to a different follow-up management.

Conclusions: BT-RADS can be considered a valid tool for neuroradiologists and radiologists even with little experience in the interpretation of patients' images during follow-up for adult primary brain tumors supporting standardized interpretation, reporting and clinical management.

[^] ORCID: 0000-0002-4067-7118.

Keywords: Inter-reader agreement; magnetic resonance imaging; imaging (MRI); gliomas; Brain Tumor Reporting and Data System (BT-RADS)

Submitted Aug 12, 2022. Accepted for publication Dec 12, 2022. Published online Jan 14, 2023.

doi: 10.21037/qims-22-850

View this article at: <https://dx.doi.org/10.21037/qims-22-850>

Introduction

Although brain metastases are the most common type of intracranial neoplasm in adults (1), gliomas represent the most frequent primary intra-axial neoplasm in adults and only 5% of patients with glioblastoma survive to 5 years (2,3).

In recent years there was a large-scale dissemination of structured reports in radiology. Following this trend, in 2018, a new system was proposed for classifying and reporting post-treatment brain tumor on magnetic resonance imaging (MRI), named as Brain Tumor Reporting and Data System (BT-RADS) (4).

Brain MRI is the primary imaging technique in brain tumor surveillance. The need for this score arises from the evidence that it is often difficult interpreting images in the follow-up of brain tumors, especially when the treatments performed can mimic disease progression (pseudoprogression) and medication may reflect a not truly tumor response (pseudoresponse) (5), leading to inconsistent and ambiguous reports. Perfusion MRI techniques can be useful in the differentiation between glioma recurrence and pseudoprogression; nevertheless, these methods are not yet widely available and requires further standardization in term of parameters and threshold differences (6). Macdonald and Response Assessment in Neuro-Oncology (RANO) criteria have been previously defined for assessing brain tumor response (7,8), and are currently suggested in the Italian oncological guidelines for the follow up of brain tumors (9). However, these criteria are not widespread in clinical practice due to their complexity, need for multiple measurements, lack of understanding by clinicians and interobserver variability (10).

For the correct evaluation of BT-RADS it is necessary to know the patient's clinical history in term of surgical history, time elapsed since radiotherapy (more or less than 90 days) and the therapeutic regimen applied (especially Avastin and steroids). In addition, MR image analysis is performed with an emphasis on fluid attenuated inversion recovery (FLAIR) and T1 post-contrast images. Based on these findings, the

BT-RADS provides a numerical score from 0 to 4 with some sub-categories as regards the score 1 (divided into 1a and 1b) and 3 (divided into 3a, 3b and 3c), each of these with simple management recommendations (4).

The prognostic role of this score has already been recently validated in another study where it was found that the score with the highest probability of worsening at subsequent follow-up was 3b and that mortality increased by 53% for each increase in the BT-RADS score during surveillance (11).

As with other widely used RADS (12,13), it is necessary to validate the BT-RADS score also by means of agreement studies. The goal of our study is to assess the interrater reliability of BT-RADS among expert radiologists and radiologists under training on a single institution cohort of brain tumor MRIs.

We present this article in accordance with the GRRAS reporting checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-22-850/rc>).

Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). An informed consent to the possible use of existing clinical data for research purposes was obtained from all patients as an institutional policy. The ethics committee of Fondazione Policlinico Universitario Campus Bio-Medico di Roma approved the study (ComEt 0050/20) and waived the informed consent based on the observational design of the study and because only secondary analysis of existing anonymized data was performed.

A retrospective study was designed with the aim of identifying contrast-enhanced MRI of adult patients on follow-up of primary brain tumor at Fondazione Policlinico Campus Bio-Medico with at least two consecutive imaging studies after surgery. The time interval taken into consideration was from January 1st 2006 to February 1st 2022. All MRI included pre-contrast FLAIR images and T1-weighted images after intravenous administration of

Table 1 Distribution of patients included in the study

Variables	Value
Number of patients (male/female)	23 (14/9)
Mean age \pm SD [range], years	54 \pm 16 [20–83]
Number of MRIs	147
Number of BT-RADS	588
Tumor types (number of patients)	
Astrocytoma grade 4 (14/23, 61%)	90/147 (62%)*
Astrocytoma grade 3 (4/23, 17%)	11/147 (7%)*
Astrocytoma grade 2 (1/23, 4%)	12/147 (8%)*
Oligodendroglioma grade 3 (3/23, 13%)	31/147 (21%)*
Oligodendroglioma grade 2 (1/23, 4%)	3/147 (2%)*

*, number of MRI assessed per single tumor histological subtype. SD, standard deviation; MRI, magnetic resonance imaging; BT-RADS, Brain Tumor Reporting and Data System.

gadolinium chelate as required for the BT-RADS evaluation. In particular, images were acquired using a 1.5 Tesla MRI system (Magnetom Symphony or Avanto B13, Siemens, Erlangen, Germany), configured with a 12-element designed Head Matrix Coil. The imaging protocol included: axial fluid attenuated inversion recovery (FLAIR: TR 8,000 ms, TE 102 ms, TI 3,650 ms, matrix 256 \times 256, FOV 26 \times 30 cm, slice thickness 3 mm), coronal turbo spin echo (TSE), T2 weighted (TR 6,380 ms, TE 105 ms, matrix 256 \times 256 FOV 26 \times 30 cm, slice thickness 3 mm) with and without fat suppression, and multiplanar (i.e., coronal, sagittal and axial) two-dimensional (2D) turbo spin echo (TSE) T1 weighted (TR 663 ms, TE 11 ms, matrix 256 \times 256, FOV 25 \times 25 cm, slice thickness 3 mm, in-plane resolution of 1 \times 1 mm²) sequences, conducted before and after gadolinium containing contrast agents (Dotarem[®], gadoterate meglumine 0.5 mmol/mL or Omniscan[®], gadodiamide 0.5 mmol/mL) intra-venous administration. The following data were collected for each patient from our electronic medical record system and PACS: sex, type of brain tumor, surgical history, end date of radiation treatment, start date of medication treatment and date of MRIs performed. Four radiologists [2 residents in radiology (M.Pi. and D.Di.), 1 general radiologist with 4 years of experience (M.Pa.) and 1 neuroradiologist with 10 years of experience (C.A.Ma.)] read and scored each study using the BT-RADS scoring tool available on <https://btrads.com/scoring/>, blinded to the MRI original report.

Prior to beginning the study, readers were provided with an online teaching file (<https://btrads.com/>) consisting of a brief background about BT-RADS and practice cases (one case for each of the seven BT-RADS categories).

Cases were accessed through any Carestream Vue PACS v. 12.2 (Carestream Health, Rochester, New York, United State of America) commonly used in our hospital.

Descriptive statistics were collated. Interobserver agreement and Fleiss' κ were calculated to assess the level of diagnostic agreement (14). The κ statistic was interpreted using six different categories: less than 0, no agreement; 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–0.99, almost perfect agreement (15).

In addition, it was assessed how many times the assignment of different scoring of BT-RADS would have led to a different patient management. In fact, the assignment to different categories affects the recommended management: i.e., continued follow-up for categories 1 and 2, decreased time interval of follow-up for categories 3a and 3b, consider change in management or decreased time interval of follow-up for category 3c and consider change in management for category 4 (4).

All the statistical analyses were conducted by using the Statistical Package for the Social Sciences software package (SPSS version 28.0.0, IBM Corp., Armonk, New York, United State of America), setting the a priori significance level at $P < 0.05$.

Results

The total number of patients included in the study was 23 (14 males and 9 females with a mean age of 54 \pm 16 years) with 147 MRIs and a total of 588 BT-RADS scores retrospectively evaluated. The two most frequent tumor types were astrocytoma grade 4 (90/147, 62%) and oligodendroglioma grade 3 (31/147, 21%); the other types of brain tumors in order of frequency were: astrocytoma grade 2 (12/147, 8%), astrocytoma grade 3 (11/147, 7%) and oligodendroglioma grade 2 (3/147, 2%) (*Table 1*). All patients with oligodendroglioma (3/23, 13% with oligodendroglioma grade 3 and 1/23, 4% with oligodendroglioma grade 2), 12/23 (52%) patients with astrocytoma grade 4, 2/23 (9%) patients with astrocytoma grade 3 and 1/23 (4%) patient with astrocytoma grade 2 were classified following the fourth edition of the World Health Organization (WHO) classification of central nervous system tumors (16). 1/23 (4%) patient with astrocytoma grade 4 and 1/23 (4%)

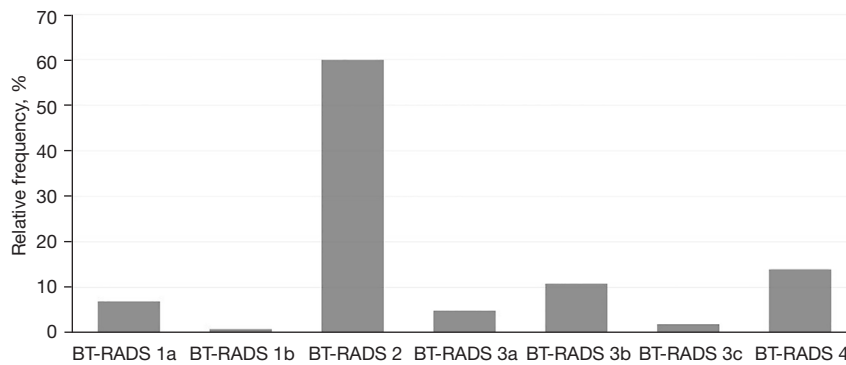


Figure 1 Column chart showing the relative frequency of the BT-RADS scores evaluated. BT-RADS, Brain Tumor Reporting and Data System.

Table 2 Interrater agreement

Rate of agreement	Percentages	k	95% CI	Level of agreement
All cases	82%	0.70	0.62–0.78	Substantial agreement
Astrocytoma grade 4	82%	0.70	0.60–0.80	Substantial agreement
Astrocytoma grade 3	61%	0.51	0.39–0.64	Moderate agreement
Astrocytoma grade 2	92%	0.78	0.33–1.00	Substantial agreement
Oligodendroglioma grade 3	90%	0.78	0.57–0.99	Substantial agreement
Oligodendroglioma grade 2	50%	0.32	0.03–0.61	Slight agreement
All cases: RAD vs. NEURORAD	91%	0.86	0.74–0.98	Almost perfect agreement
All cases: RES vs. NEURORAD	80%	0.66	0.57–0.76	Substantial agreement

Number of readers in agreement: 4 radiologists agree [103/147 (70%)]; 3 radiologists agree [25/147 (17%)]; 2 radiologists agree [19/147 (13%)]; no agree (0). RAD, general radiologist; RES, radiology residents; NEURORAD, neuroradiologist; k, Fleiss' kappa; CI, confidence interval.

patient with astrocytoma grade 3 were classified following the revised fourth edition of the WHO classification of central nervous system tumors (17). 1/23 (4%) patient with astrocytoma grade 4 and 1/23 (4%) patient with astrocytoma grade 3 were classified following the fifth edition of the WHO classification of central nervous system tumors (18).

BT-RADS score categories were distributed as it follows with most reported studies into category 2 (Figure 1): score 1a 39/588 (7%), score 1b 8/588 (1%), score 2 353/588 (60%), score 3a 28/588 (5%), score 3b 65/588 (11%), score 3c 13/588 (2%) and score 4 82/588 (14%).

The overall agreement rate for all 4 radiologists was 82% with a Fleiss' k of 0.70 (substantial agreement). The overall agreement rate between general radiologist and neuroradiologist was 91% with a Fleiss' k of 0.86 (almost

perfect agreement). The overall agreement rate between 2 radiology residents and neuroradiologist was 80% with a Fleiss' k of 0.66 (substantial agreement). When reviewing tumor subtypes, astrocytoma grade 3 (k: 0.51, moderate agreement) and oligodendroglioma grade 2 (k: 0.32, fair agreement) showed a poor agreement while higher values of agreement were found for astrocytoma grade 4 (k: 0.70, substantial agreement), astrocytoma grade 2 (k: 0.78, substantial agreement) and oligodendroglioma grade 3 (k: 0.78, substantial agreement). All the radiologists agreed on BT-RADS assignment in 103/147 patients (70%), three radiologists agreed in 25/147 (17%) and two radiologists agree in 19/147 (13%). In no cases there was a complete disagreement among the readers (Table 2). In 27/147 cases (18%) the discrepancy in the estimated BT-RADS would have led to a different follow-up management.

Figures 2,3 were two exemplary cases that demonstrate discrepancies between our raters.

Discussion

To our knowledge, this study represents the first quantifiable analysis of blinded interrater reliability for utilizing BT-RADS to score MRIs in a European cohort of brain tumor patients.

The BT-RADS score can potentially apply to any brain MRI scan during follow up of patients with primary brain tumor, as it does not require advanced MR techniques such as brain perfusion or spectroscopy.

In other realities, the implementation of BT-RADS has led to an improvement in the consistency and completeness of radiology reports and had improved communication between radiologists and referring physicians, decreasing ambiguity in the images interpretation (19,20).

In our population the BT-RADS score proved to have a good interrater reliability (substantial agreement) with a high percentage of agreement among expert radiologists and radiologists under training. In fact, we showed that BT-RADS reporting agreement was not significantly affected by physician' experience (Fleiss' $k > 0.60$), when MRI scans were evaluated by radiologists with different levels of expertise, such as a general radiologist and radiologists in training. Indeed, we found almost perfect agreement between the general radiologist and the neuroradiologist and substantial agreement between 2 radiology residents and the neuroradiologist. It is important to note that a learning curve may exist when using new templates/structured reporting; therefore, the agreement could be even greater after the systematic introduction of this score into daily clinical practice. As it has been proposed for other Reporting and Data System (21), the adoption of the BT-RADS could represent an effective strategy to facilitate and speed up the learning curve of radiologists in the evaluation of MRI scans of patients treated for brain tumor. Moreover, the possibility of having easy access to an online tool to calculate the BT-RADS improves the learning of the score itself (22).

In 18% of MRIs the assignment of BT-RADS would have led to a different follow-up management; we hypothesize that even this figure can be further improved with the institutional systematic use of the score.

We observed poor agreement in the sub-analysis of some tumor types, including astrocytoma grade 3 and

oligodendroglioma grade 2 (11/147 cases for the first, 3/147 cases for the latter). On the other hand, a strong agreement was found for astrocytoma grade 4 that represent the commonest malignant primary brain tumor in adults (3). While the reason for this discrepancy is uncertain, it may represent challenges in evaluating tumors that have no or little contrast enhancement, as the criteria depend heavily on post-contrast imaging. In addition, the small number of cases of low-grade tumors does not allow a definitive agreement analysis in these cancer subtypes.

It is worth to mention that our study results should be taken with caution due to some limitations. First, the retrospective design of the study did not allow to have clear access to all the clinical information of the selected patients. Second, the small number of patients can be considered a limitation of the study; nevertheless, since BT-RADS is a score whose purpose is to perform patient surveillance (and not to make a diagnosis), the value of the study should rely on the total number of follow-up MRI examinations ($n=147$) assessed for longitudinal evaluation rather than the number of patients ($n=23$). Third, the evaluation of tumor subtypes (excluding grade IV astrocytoma) could have been affected by the fact that we considered patients histologically assessed over a wide time interval of about 15 years; that resulted in various definitions of tumor subtypes according to the different versions of the WHO classification. Despite this issue on histological definition of tumors, we do not consider it as a major bias because the image evaluation and scoring system is independent on the histological type. Fourth, the low number of complex cases as demonstrated by the high relative frequency of BT-RADS 2 might have simplified assessment, increasing the agreement rate. However, this does reflect the overall distribution of tumors in imaging practice. Fifth, the analysis of cases from a single institution database does not allow to make any conclusion on reproducibility and generalizability in different clinical and technological settings.

Further research will focus on assessing multi-institutional BT-RADS interrater reliability in a prospective way, also evaluating a larger number of cancer subtypes according to the last version of WHO brain tumor classification available.

In summary, we show that BT-RADS can be considered a valid tool for neuroradiologists and radiologists even with little experience in the interpretation of patients' images during follow-up for adult primary brain tumors supporting standardized interpretation, reporting and clinical management.

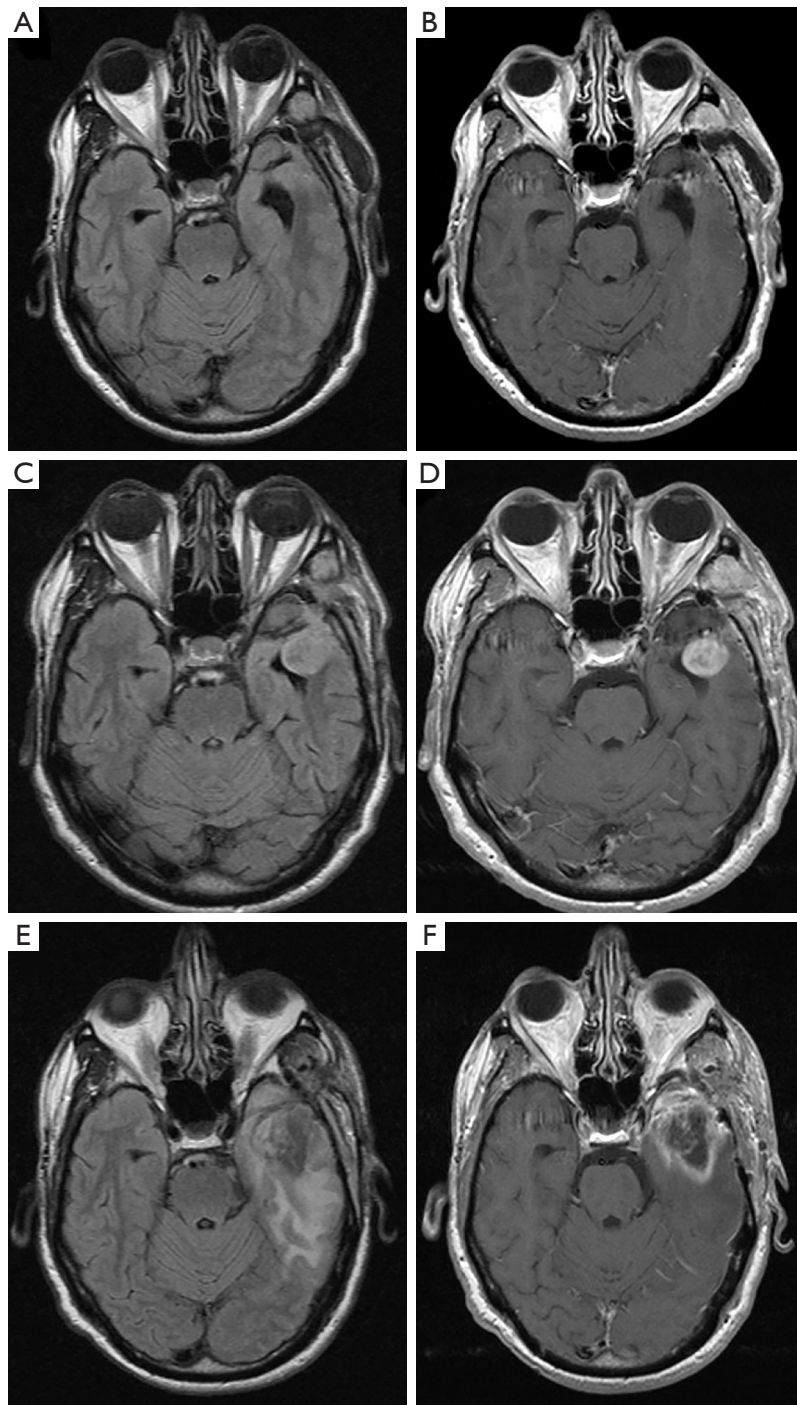


Figure 2 Astrocytoma grade 4 of the left temporal lobe surgically treated and subsequently underwent radiation therapy. (A,C,E) Axial FLAIR images; (B,D,F) axial T1 contrast-enhanced images. (A,B) MRI performed after 3 months from the surgery with no obvious signs of residual disease. (C,D) Follow up MRI performed after 1 months from A and B showing a nodular mass with contrast enhancement in the left temporal pole, highly suspected for disease recurrence. Half of the raters assigned a BT-RADS category 3c. However, given the worsening >25% in FLAIR and enhanced images, BT-RADS category 4 is more appropriate. (E,F) Follow up MRI after 2 months (<90 days) from the radiation therapy showing a worsening in FLAIR and enhanced images still to be interpreted as related to the treatment. All readers assigned a BT-RADS category of 3a. FLAIR, fluid attenuated inversion recovery; MRI, magnetic resonance imaging; BT-RADS, Brain Tumor Reporting and Data System.

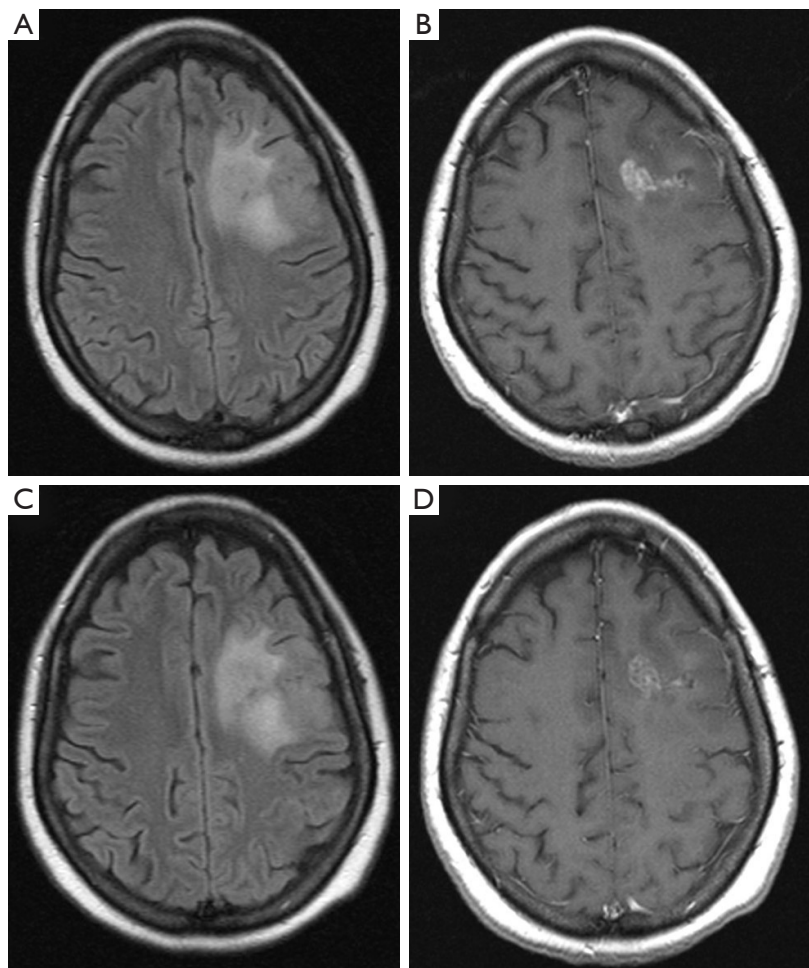


Figure 3 Oligodendroglioma grade 2 of the left frontal lobe surgically treated and subsequently underwent radiation therapy. (A,C) Axial FLAIR images; (B,D) axial T1 contrast-enhanced images. (A,B) MRI performed after 7 months from the surgery and after 3 months from the radiation therapy with a residual area of contrast enhancement and FLAIR hyperintensity. (C,D) Follow up MRI performed after 3 months from (A) and (B) showing a mild decrease in the FLAIR hyperintensity and contrast-enhancement. One rater assigned a BT-RADS category 2 as for stable disease. However, given the improvement in FLAIR and enhanced images, BT-RADS 1a is more appropriate. FLAIR, fluid attenuated inversion recovery; MRI, magnetic resonance imaging; BT-RADS, Brain Tumor Reporting and Data System.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: With the arrangement by the Guest Editors and the editorial office, this article has been reviewed by external peers.

Reporting Checklist: The authors have completed the GRRAS reporting checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-22-850/rc>

<https://qims.amegroups.com/article/view/10.21037/qims-22-850/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-22-850/coif>). The special issue “Imaging of Aging and Age-Related Disorders” was commissioned by the editorial office without any funding or sponsorship. CAM served as the unpaid Guest Editor of the issue and serves as an unpaid editorial board member of *Quantitative Imaging in Medicine and Surgery*. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional ethics board of Fondazione Policlinico Universitario Campus Bio-Medico di Roma (ComEt 0050/20) and individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Sacks P, Rahman M. Epidemiology of Brain Metastases. *Neurosurg Clin N Am* 2020;31:481-8.
- Yin L, Cheng L, Wang F, Zhu X, Hua Y, He W. Application of intraoperative B-mode ultrasound and shear wave elastography for glioma grading. *Quant Imaging Med Surg* 2021;11:2733-43.
- McNeill KA. Epidemiology of Brain Tumors. *Neurol Clin* 2016;34:981-98.
- Weinberg BD, Gore A, Shu HG, Olson JJ, Duszak R, Voloschin AD, Hoch MJ. Management-Based Structured Reporting of Posttreatment Glioma Response With the Brain Tumor Reporting and Data System. *J Am Coll Radiol* 2018;15:767-71.
- Quant EC, Wen PY. Response assessment in neuro-oncology. *Curr Oncol Rep* 2011;13:50-6.
- Zhang J, Wang Y, Wang Y, Xiao H, Chen X, Lei Y, Feng Z, Ma X, Ma L. Perfusion magnetic resonance imaging in the differentiation between glioma recurrence and pseudoprogression: a systematic review, meta-analysis and meta-regression. *Quant Imaging Med Surg* 2022;12:4805-22.
- Huang RY, Wen PY. Response Assessment in Neuro-Oncology Criteria and Clinical Endpoints. *Magn Reson Imaging Clin N Am* 2016;24:705-18.
- Macdonald DR, Cascino TL, Schold SC Jr, Cairncross JG. Response criteria for phase II studies of supratentorial malignant glioma. *J Clin Oncol* 1990;8:1277-80.
- Linee Guida AIOM Neoplasie Cerebrali. Available online: https://snlg.iss.it/wp-content/uploads/2021/10/LG_266_neoplasie_cerebrali_agg2021.pdf (accessed 2022-05-11).
- Vos MJ, Uitdehaag BM, Barkhof F, Heimans JJ, Baayen HC, Boogerd W, Castelijns JA, Elkhuisen PH, Postma TJ. Interobserver variability in the radiological assessment of response to chemotherapy in glioma. *Neurology* 2003;60:826-30.
- Kim S, Hoch MJ, Peng L, Somasundaram A, Chen Z, Weinberg BD. A brain tumor reporting and data system to optimize imaging surveillance and prognostication in high-grade gliomas. *J Neuroimaging* 2022;32:1185-92.
- Hsu D, Rath TJ, Branstetter BF, Anzai Y, Phillips CD, Juliano AF, Mosier KM, Bazylewicz MP, Poliashenko SM, Kulzer MH, Rhyner PA, Risk B, Wiggins RH, Aiken AH. Interrater Reliability of NI-RADS on Posttreatment PET/Contrast-enhanced CT Scans in Head and Neck Squamous Cell Carcinoma. *Radiol Imaging Cancer* 2021;3:e200131.
- Del Giudice F, Pecoraro M, Vargas HA, Cipollari S, De Berardinis E, Bicchetti M, Chung BI, Catalano C, Narumi Y, Catto JWF, Panebianco V. Systematic Review and Meta-Analysis of Vesical Imaging-Reporting and Data System (VI-RADS) Inter-Observer Reliability: An Added Value for Muscle Invasive Bladder Cancer Detection. *Cancers (Basel)* 2020.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:360-3.
- Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, Scheithauer BW, Kleihues P. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* 2007;114:97-109.
- Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol* 2016;131:803-20.
- Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, Hawkins C, Ng HK, Pfister SM, Reifenberger G, Soffietti R, von Deimling A, Ellison DW. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol* 2021;23:1231-51.

19. Gore A, Hoch MJ, Shu HG, Olson JJ, Voloschin AD, Weinberg BD. Institutional Implementation of a Structured Reporting System: Our Experience with the Brain Tumor Reporting and Data System. *Acad Radiol* 2019;26:974-80.
20. Zhang JY, Weinberg BD, Hu R, Saindane A, Mullins M, Allen J, Hoch MJ. Quantitative Improvement in Brain Tumor MRI Through Structured Reporting (BT-RADS). *Acad Radiol* 2020;27:780-4.
21. Bellini D, Panvini N, Rengo M, Vicini S, Lichtner M, Tieghi T, Ippoliti D, Giulio F, Orlando E, Iozzino M, Ciolfi MG, Montechiarello S, d'Ambrosio U, d'Adamo E, Gambaretto C, Panno S, Caldon V, Ambrogi C, Carbone I. Diagnostic accuracy and interobserver variability of CO-RADS in patients with suspected coronavirus disease-2019: a multireader validation study. *Eur Radiol* 2021;31:1932-40.
22. Kim S, Hoch MJ, Cooper ME, Gore A, Weinberg BD. Using a Website to Teach a Structured Reporting System, the Brain Tumor Reporting and Data System. *Curr Probl Diagn Radiol* 2021;50:356-61.

Cite this article as: Parillo M, Mallio CA, Pileri M, Dirawe D, Romano A, Bozzao A, Weinberg B, Quattrocchi CC. Interrater reliability of Brain Tumor Reporting and Data System (BT-RADS) in the follow up of adult primary brain tumors: a single institution experience in Italy. *Quant Imaging Med Surg* 2023;13(11):7423-7431. doi: 10.21037/qims-22-850