

UCLA

UCLA Previously Published Works

Title

On the promise of artificial intelligence for standardizing radiographic response assessment in gliomas

Permalink

<https://escholarship.org/uc/item/3d7289nc>

Journal

Neuro-Oncology, 21(11)

ISSN

1522-8517

Author

Ellingson, Benjamin M

Publication Date

2019-11-04

DOI

10.1093/neuonc/noz162

Peer reviewed

**On the promise of artificial intelligence for standardizing radiographic response  
assessment in gliomas**

Benjamin M. Ellingson, Ph.D.<sup>1</sup>

<sup>1</sup> UCLA Brain Tumor Imaging Laboratory, Center for Computer Vision and Imaging  
Biomarkers, Department of Radiological Sciences, David Geffen School of Medicine,  
University of California Los Angeles, Los Angeles, CA 90024.

*(Editorial on*

*"Automatic assessment of glioma burden: A deep learning algorithm for fully automated  
volumetric and bi-dimensional measurement")*

Address Correspondence To:

Benjamin M. Ellingson, Ph.D.  
Professor of Radiology, Biomedical Physics, Bioengineering, and Psychiatry  
Departments of Radiological Sciences and Psychiatry  
David Geffen School of Medicine  
University of California Los Angeles  
924 Westwood Blvd, Suite 615  
Los Angeles, CA 90024  
bellingson@mednet.ucla.edu

Downloaded from <https://academic.oup.com/neuro-oncology/advance-article-abstract/doi/10.1093/neuonc/nwz192/5581877> by UCLA Biomedical Library Serials user on 01 October 2019

Identification and delineation of tumor burden on MR images is a critical, yet challenging, function for clinical care and neuro-oncology research. Changes in tumor burden on structural MRI sequences are used as surrogates for efficacy throughout the entire drug development pipeline, from initial preclinical evaluations through clinical tests in human patients. *DeepNeuro*, the algorithm recently published by Chang *et al.*<sup>1</sup>, and other automated approaches to tumor segmentation (see reviews by Gordillo *et al.*<sup>2</sup>, Iyji *et al.*<sup>3</sup>, and Bauer *et al.*<sup>4</sup>), combined with standardization of image acquisition and knowledgeable interpretation of radiographic changes, holds great promise for accurate, repeatable, efficient, and inexpensive tumor segmentation.

Manual or semi-automated quantification of tumor size by experts using commercial or academic tools is highly variable across platforms<sup>5</sup> and is extremely time consuming, labor intensive, inefficient, and expensive. Expert neuroradiologists in large clinical trials average more than 40% adjudication rate<sup>6,7</sup>, which is consistent with adjudication or discrepancy rates between radiologists in other tumor types and between central review compared with local site call of progression (~30-40%)<sup>8</sup>. Difficulty in determining changes in T2 hyperintense lesions on RANO was often determined to be the largest reason for adjudication, while changes in contrast enhancement appeared less ambiguous, but still relatively difficult. While some investigators have suggested volumetric segmentation on T1 subtraction maps may be a way to replace blinded, adjudicated central reads<sup>9</sup>, there remains significant questions as to the reproducibility of simple segmentation approaches in the presence of artifacts and generalization of these threshold-based segmentation techniques for all relevant components of the tumor (e.g. enhancing, necrotic, and non-enhancing disease). Automated segmentation and quantitative

response assessment using a large set of post-therapeutic data, as demonstrated by *DeepNeuro*<sup>1</sup>, has significant potential to reduce these adjudication rates while allowing investigators and regulators to evaluate potential changes in various aspects of the tumor biology.

In addition to standard response assessment, clearly defining regions of interest for both T2 hyperintense lesions (e.g. edema and/or non-enhancing tumor) and contrast enhancing lesions on post-contrast T1-weighted images are critical for obtaining advanced imaging measurements (diffusion, perfusion, PET, etc.) within areas of active disease. Thus, tumor segmentation itself adds another level of uncertainty beyond the inherent variability in advanced image acquisition and post-processing. The use of an automated and standardized tumor segmentation algorithm like *DeepNeuro* could conceivably reduce variability in advanced imaging analyses and increase repeatability of results across centers and studies.

Performance of sophisticated algorithms like *DeepNeuro* are going to greatly improve through the adaptation and implementation of the standardized brain tumor imaging protocol (BTIP)<sup>11</sup>. Introduced in 2015 as a result of a joint meeting between the FDA, NCI, clinical scientists, imaging experts, pharmaceutical and biotech companies, clinical trial cooperative groups, and patient advocacy groups, BTIP is compliant with ACRIN and EORTC guidelines and includes the minimum recommended sequences for response assessment using 1.5T and 3T MRI scanners. Since most machine learning algorithms utilize public datasets obtained from individual institutions or trials that may not be BTIP compliant, they are innately designed to generalize across a wide range of acquisition parameters and machine settings. By streamlining and focusing the range of MR parameters that are acceptable for use in clinical trials, performance of automated algorithms will theoretically improve as trials comply with BTIP and adapt to less variability in imaging characteristics.

"We are drowning in data but starved for wisdom."

~ Arianna Huffington, Founder of Huffington Post re: Artificial Intelligence

More than 200 scientific abstracts, plenary talks, and presentations were devoted to artificial intelligence or machine learning at the 2018 Radiologic Society of North America (RSNA) annual meeting last year. In radiology, artificial intelligence and "Deep Learning" approaches are changing the way we schedule patients, acquire and post-process images, measure tumor size, dictate or interpret changes, and charge for radiological services. While *DeepNeuro* and other techniques with automatic pipelines hold promise of reducing ambiguity and variability in tumor segmentation, these algorithms are largely based on "expert reads", which may be problematic as experts may not be properly vetted or trained, may have difficulty accurately defining tumor boundaries in diffuse gliomas, and/or they may not utilize all approaches we know to be useful for defining active disease (e.g. T1 subtraction maps, subtle signal intensity changes on T2-weighted images, architectural disruption at the gray-white matter boundary, etc.). Additionally, most of the available automated algorithms have been trained entirely on treatment naïve tumors<sup>5</sup>, which is a much simpler problem (i.e. no post-surgical or post-therapeutic changes) with no tangible clinical impact. It is more meaningful, and admittedly more difficult, to try and differentiate changes in tumor during the post-therapeutic setting or during treatment with experimental therapeutics. While *DeepNeuro* may not have not yet undergone extensive testing under various therapeutic contexts (e.g. anti-angiogenic or immunotherapies), or in challenging tumor types (e.g. gliomatosis cerebri, mixed grade tumors, etc.), their attempt to adapt their algorithm to perform an automated response assessment during treatment within a clinical trial is a strong, noteworthy step in the right direction. Thus, while

promising, artificial intelligence is inherently only as good as the knowledge we are able to provide. These algorithms currently cannot generate new wisdom into the intrinsic biology of brain tumors, nor can they begin to unravel the true extent of tumor infiltration into the brain. Perhaps someday, with the addition of autopsy data, this later goal could become a reality. However, just like humans, these new A.I. algorithms have the ability to adapt, adjust, learn, and “self-correct” to improve performance based on new or improved knowledge.

We are at the dawn of a renaissance in imaging technology in neuro-oncology. Artificial intelligence-based techniques and approaches for automated response assessment are going to increase in sophistication and complexity in the years and decades to come. Investigators and regulators should be acutely aware of the limitations of this technology, providing wisdom and context to further guide the development and implementation in both clinical trials and clinical care for the benefit of brain tumor patients.

Accepted Manuscript

## References

1. Chang K, Beers AL, Bai HX, et al. Automatic assessment of glioma burden: A deep learning algorithm for fully automated volumetric and bi-dimensional measurement. *Neuro Oncol*. 2019. doi.org/10.1093/neuonc/noz106
2. Gordillo N, Montseny E, Sobrevilla P. State of the art survey on MRI brain tumor segmentation. *Magn Reson Imaging*. 2013; 31(8):1426-1438.
3. İşin A, Direkçoglu C, Sah M. Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. *Procedia Computer Science*. 2016; 102:317-324.
4. Bauer S, Wiest R, Nolte LP, Reyes M. A survey of MRI-based medical image analysis for brain tumor studies. *Phys Med Biol*. 2013; 58(13):R97-129.
5. Menze BH, Jakab A, Bauer S, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans Med Imaging*. 2015; 34(10):1993-2024.
6. Ford RR, O'Neal M, Moskowitz SC, Fraunberger J. Adjudication rates between readers in blinded independent central review of oncology studies. *J Clin Trials*. 2016; 6(5):289.
7. Boxerman JL, Zhang Z, Saffriel Y, et al. Early post-bevacizumab progression on contrast-enhanced MRI as a prognostic marker for overall survival in recurrent glioblastoma: results from the ACRIN 6677/RTOG 0625 Central Reader Study. *Neuro Oncol*. 2013; 15(7):945-954.
8. Pope WB, Hessel C. Response assessment in neuro oncology criteria: implementation challenges in multicenter neuro-oncology trials. *AJNR Am J Neuroradiol*. 2011; 32(5):794-797.
9. Dodd LE, Korn EL, Freidlin B, et al. Blinded independent central review of progression-free survival in phase III clinical trials: important design element or unnecessary expense? *J Clin Oncol*. 2008; 26(22):3791-3796.
10. Schmainda KM, Prah MA, Zhang Z, et al. Quantitative Delta T1 (dT1) as a Replacement for Adjudicated Central Reader Analysis of Contrast-Enhancing Tumor Burden: A Subanalysis of the American College of Radiology Imaging Network 6677 Radiation Therapy Oncology Group 0625 Multicenter Brain Tumor Trial. *AJNR Am J Neuroradiol*. 2019; 40(7):1132-1139.
11. Ellingson BM, Bendixus M, Boxerman J, et al. Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials. *Neuro Oncol*. 2015; 17(9):1188-1198.