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
2019-11-01

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. 1

1 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 90095

(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 Sciences, Bioengineering, and Psychiatry
Departments of Radiology, Neuroscience, 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

© The Author(s) 2019. Published by Oxford University Press on behalf of the Society for Neuro-Oncology. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
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 surrogate endpoints for efficacy throughout the entire drug development pipeline, from initial preclinical evaluations through clinical trials in human patients. DeepSiac, the algorithm recently published by Chang et al., and other automated approaches to tumor segmentation (see reviews by Geroldi et al.6, Liu et al., and Banner et al.5), combined with standardization of image acquisition and automated interpretation of radiographic changes, holds great promise for accurate, reproducible, efficient, and inexpensive tumor segmentation.

Manual or semi-automated quantification of tumor size by experts using commercial or academic tools is highly variable across institutions and is extremely time consuming, labor intensive, inefficient, and expensive. Even neuroradiologists in large clinical trials average more than 40% adjudication rates, which is consistent with adjudication or discrepancy rates between radiologists in some large, prospective, and central review compared with local site calls of progression.14, 15, 16. Difficulty in determining changes in T2 hyperintense lesions on RANO was also identified to be the largest reason for adjudication, while changes in contrast enhancement appeared less ambiguous, but still relatively difficult. While some investigators have attempted volumetric segmentation on T1 subtraction maps as a way to replace blinded, adjudicated central reads, 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 posttherapeutic data, as demonstrated by DeepNeuro\textsuperscript{1}, has significant potential to reduce those 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 hypointense lesions (e.g. edema and/or necrosis) and contrast-enhancing lesions on post-contrast T1-weighted images is critical for obtaining advanced imaging assessments (diffusion, perfusion, PET, etc.) within areas of active disease. Thus, manual segmentation itself adds another level of uncertainty beyond the inherent variability in actual image acquisition and post-processing. The use of an automated and standardized tumor-segmentation algorithm like DeepNeuro could conceivably reduce variability in automated imaging analyses and increase repeatability of results across centers and studies.

Performance of sophisticated algorithms like DeepNeuro are going to greatly improve through the adoption and implementation of the standardized brain tumor imaging protocol (STBP)\textsuperscript{2}. Introduced in 2013 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, STBP is compliant with ACRIN and EORTC guidelines and includes an expanded 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 STBP compliant, they are inherently designed to generalize across a wide range of acquisition parameters and machine settings. By uncalibrating 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 STBP 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 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, those algorithms are largely based on “expert rules”, which may be problematic as experts may not be properly vetted or trained, may have difficulty accurately defining tumor boundaries or edge contours, and/or may not utilize all approaches we know to be useful for defining active disease (e.g. TI subtraction maps, subtle signal intensity changes on T2 weighted images, architectural disruption at the gray-white matter boundary, etc.). Additionally, many of the available automated algorithms have been trained entirely on training set cases (e.g. which is a much simpler problem (i.e. no post-surgical or post-therapeutic changes) with 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. gliomas of cerebrum, mixed grade tumors, etc.), their attempt to adapt their algorithms 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 latter goal could become a reality. However, just like humans, these new A.I. algorithms have the ability to adapt, improve, and “self-learn” to improve performance based on new or improved knowledge.

We are at the dawn of a renaissance in imaging technology in neurology. 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.
References


