Manual segmentation

Precise volumetric assessment of brain tumors is relevant for treatment planning and monitoring. However, manual segmentations are time-consuming and impeded by intra- and inter rater variabilities.

To investigate the performance of a deep learning model (DLM) to automatically detect and segment primary central nervous system lymphoma (PCNSL) on clinical MRI.

Study type: Retrospective.

Population: Sixty-nine scans (at initial and/or follow-up imaging) from 43 patients with PCNSL referred for clinical MRI tumor assessment.

Field strength/sequence: T1 -/T2 -weighted, T1 -weighted contrast-enhanced (T1 CE), and FLAIR at 1.0, 1.5, and 3.0T from different vendors and study centers.

Assessment: Fully automated voxelwise segmentation of tumor components was performed using a 3D convolutional neural network (DeepMedic) trained on gliomas (n = 220). DLM segmentations were compared to manual segmentations performed in a 3D voxelwise manner by two readers (radiologist and neurosurgeon; consensus reading) from T1 CE and FLAIR, which served as the reference standard.

Statistical tests: Dice similarity coefficient (DSC) for comparison of spatial overlap with the reference standard, Pearson's correlation coefficient ® to assess the relationship between volumetric measurements of segmentations, and Wilcoxon rank-sum test for comparison of DSCs obtained in initial and follow-up imaging.

Results: The DLM detected 66 of 69 PCNSL, representing a sensitivity of 95.7%. Compared to the reference standard, DLM achieved good spatial overlap for total tumor volume (TTV, union of tumor volume in T1 CE and FLAIR; average size 77.16 ± 62.4 cm3 , median DSC: 0.76) and tumor core (contrast enhancing tumor in T1 CE; average size: 11.67 ± 13.88 cm3 , median DSC: 0.73). High volumetric correlation between automated and manual segmentations was observed (TTV: r = 0.88, P < 0.0001; core: r = 0.86, P < 0.0001). Performance of automated segmentations was comparable between pretreatment and follow-up scans without significant differences (TTV: P = 0.242, core: P = 0.177).

Data conclusion: In clinical MRI scans, a DLM initially trained on gliomas provides segmentation of PCNSL comparable to manual segmentation, despite its complex and multifaceted appearance. Segmentation performance was high in both initial and follow-up scans, suggesting its potential for application in longitudinal tumor imaging.

Level of evidence: 3 TECHNICAL EFFICACY STAGE: 2 [1)].

Manual segmentation refers to the process whereby an expert transcriber segments and labels a file by hand, referring only to the spectrogram, waveform, area…

When manual segmentation is used, a low interobserver agreement in the assessment of tumor resection rates on magnetic resonance imaging (MRI) is described.

This applies particularly to post-operative tumor volume and residual tumor volume [2].

1)
Pennig L, Hoyer UCI, Goertz L, et al. Primary Central Nervous System Lymphoma: Clinical Evaluation of Automated Segmentation on Multiparametric MRI Using Deep Learning [published online ahead of print, 2020 Jul 13]. J Magn Reson Imaging. 2020;e27288. doi:10.1002/jmri.27288
2)
Kubben PL, Postma AA, Kessels AG, van Overbeeke JJ, van Santbrink H (2010) Intraobserver and interobserver agreement in volumetric assessment of glioblastoma multiforme resection. Neurosurgery 67:1329–1334. https://doi.org/10.1227/NEU. 0b013e3181efbb08