Clinical evaluation of atlas and deep learning-based automatic contouring of multiple organs at risk and clinical target volumes for breast cancer

Min Seo Choi, Byeong Su Choi, Seung Yeun Chung, Nalee Kim, Jaehee Chun, Yong Bae Kim, Jee Suk Chang, Jin Sung Kim

Department of Radiation Oncology, Yonsei University College of Medicine, Seoul, Korea

Purpose: Advancements of radiotherapy developments for breast cancer may render more precise and conformal radiotherapy, designed to target the tumor and avoid normal organs on a much more individualized level. Manual segmentation is the gold-standard in the current radiotherapy planning which, however, is a time-consuming process and prone to inter-observer variation. Given this issue, interest for auto-segmentation has been rising. The aim of our study is to determine the clinical feasibility of auto-segmentation methods for target and normal organs and specifically evaluate the feasibility of a deep-learning-based approach compared to the commercially released atlas-based segmentation solutions.

Materials and Methods: Contrast-enhanced planning CT data from 54 patients with breast cancer who underwent breast-conservation surgery was used in this study. Contours of target volumes, normal organs, and heart sub-structures were generated by atlas-based auto-segmentation (ABAS) software solutions (MIM and Mirada) with 35 atlas library subjects and deep learning-based auto-segmentation based on a fully convolutional DenseNet (FCDN) with 35 training sets. The accuracy of segmentation was assessed on 14 test patients using the Dice similarity coefficient (DSC) with expert-delineated manual contours as the ground truth.

Results: Compared to ABAS, the proposed FCDN model yielded more consistent results and the highest average DSC in the majority of the structures, especially the smaller CTVs where the difference of DSC was the highest: on average, FCDN produced 0.77, MIM and Mirada produced 0.68 and 0.72. In the OARs, FCDN produced average DSCs of 0.86 whereas MIM and Mirada produced 0.84 and 0.86, respectively. Lastly, in the heart and its substructures, apart from the coronary arteries, the results of FCDN (DSC: 0.72) and ABAS (MIM:0.63, Mirada:0.62) were comparable.

Conclusion: In summary, we assessed the clinical feasibility of the ABAS of MIM and Mirada and DLBAS using FCDN algorithms for the segmentation of target volumes and OARs including heart substructures. Compared to the ABAS, DLBAS using FCDN algorithm generated a more consistent and robust performance across most structures. As a preclinical study, we have confirmed the plausibility for clinical implementations of these segmentation solutions. The clinical utilization of auto-segmentation using either ABAS or DLBAS for optimal therapeutic ratio in an individualized RT plan for breast cancer, further multi-institutional collaborations is still needed.

Keyword: Deep learning-based segmentation, Atlas-based segmentation, Auto-segmentation