Multi-centre evaluation of atlas-based and deep learning contouring using a modified Turing Test


1Mirada Medical Limited, Oxford, United Kingdom. 2Radiotherapy Centre, Clinical Oncology, Cardiff, United Kingdom. 3South East Wales Cancer Centre, Clinical Oncology, Swansea, United Kingdom. 4University Medical Center Groningen, Department of Radiation Oncology, Groningen, The Netherlands. 5Radioboundaries Medical Center, Department of Radiation Oncology, Alphen, The Netherlands. 6Leeds Cancer Centre, Leeds United Kingdom. 7Imperial College Healthcare NHS Trust, Radiotherapy Department, London, United Kingdom. 8MAASTRO Clinic, Department of Radiation Oncology, Maastricht, The Netherlands.

Objective

While quantitative assessment of autocontouring quality is useful, frequently used measures do not necessarily indicate clinical acceptability or benefit. In contrast, clinical based assessment metrics, such as time saved with autocontouring or subjective evaluations, are both time consuming to perform and difficult to implement in a multi-centre evaluation.

Inspiration is taken from the Artificial Intelligence community to propose an assessment method based on the “Turing Test”. The objective of this study was to perform a multi-centre evaluation of two autocontouring methods using this approach.

Materials and Methods

A website was set up to facilitate multi-centre comparison, showing images and contours in a blinded fashion.

Use the QRcode to try it for yourself at http://www.autocontouring.com

Results

Thoracic segmentation

Classification of contours

Acceptance of contours (%)

Prostate segmentation

Classification of contours

Acceptance of contours (%)

Conclusions

The Turing Test style assessment method provided an easy way to perform web-based multicentre validation of autocontouring.

This study found that autocontours may be confused with clinical ones, when reviewed blindly. DLC showed increased clinical acceptance for prostate OAR contouring compared to atlas contouring. For thoracic imaging, DLC contours were accepted at a similar rate to clinical ones.