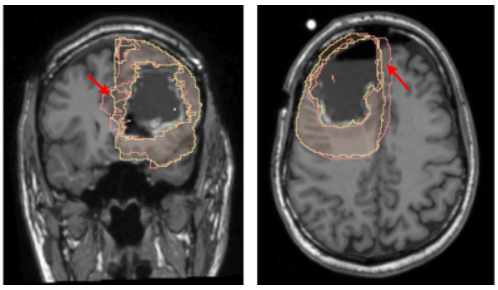
A distribution of auto-generated simulation of segmentation provides a richer comparison basis to evaluate expert human-segmentation than currently used binary overlap metrics

Automating QA for Radiation Therapy

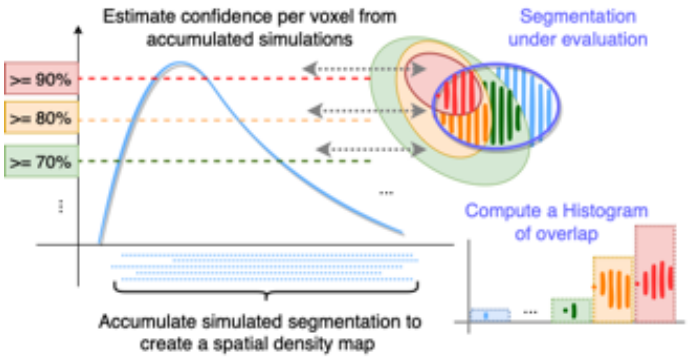
Amith Kamath[a], Robert Munger[a], Robert Poel[b], Elias Rufenacht[a], Alain Jungo[a], Ekin Ermis[a], Jonas Willmann[c], Nicolaus Andratschke[c], Mauricio Reyes[a]

1 Why QA?



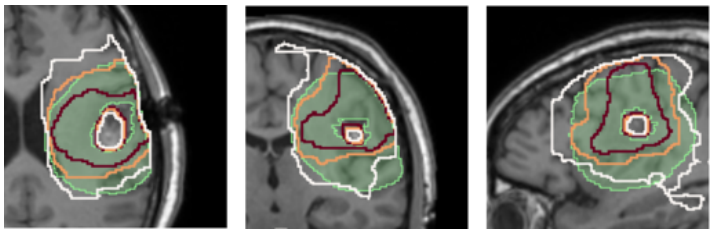
Even among experts, there is significant variability in contouring Tumors (CTV shown above) and Organs at Risk.
Hence, Radio-oncologists spend considerable time in manual QA: this cannot scale for large studies, and calls for automation!

2 Simulate realistic variability

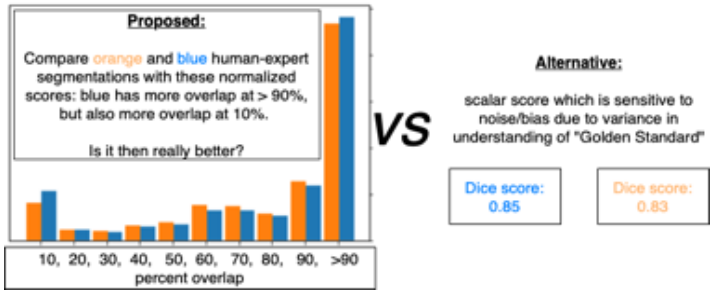


Use Deep Networks to simulate variability among human-expert segmentation: provide ability to generate more detailed evaluations.

3 Early Unpublished Results



Contours (for CTV) with several levels of confidence (red: $> 90\%$, orange: $> 50\%$, white: $> 10\%$) generated to compare with the test-segmentation (in green).



Information-rich metrics, rather than scalars like Dice coefficient or Hausdorff distance could aid in better and automated QA. Using these ideas, we intend to build a tool for segmentation Quality Assurance for Radiotherapy.

Affiliations [a]: ARTORG Center, Uni BE, [b]: Inselspital, Uni BE, [c]: Universitatspital Zurich, UZH

Abbreviations
CTV: Clinical Target Volume
QA: Quality Assurance

krebsliga schweiz
ligue suisse contre le cancer
lega svizzera contro il cancro
USZ Universitäts Spital Zürich
UNIVERSITÄT BERN