JULY 19–22
2026
VANCOUVER, BC
JOINT AAPM | COMP MEETING

LMU KLINIKUM

Automatic Tumor Segmentation for Response Assessment: RECIST Measurements as Visual Prompts for nnInteractive Model

Denis Dudas[1,3], Michael Ingrisch[1,3], Tom Julius Blöcker[1], Miguel A. Palacios[2], Chukwuka Eze[1], Claus Belka[1,3,4], Christopher Kurz[1], Guillaume Landry[1,3]

[1] LMU University Hospital Munich, Germany | [2] Amsterdam UMC, Netherlands | [3] BZKF Munich | [4] DKTK Munich

ABSTRACT

Background:
- RECIST measurements suffer from inter-observer variability and provide limited morphological information; volumetric segmentation offers more comprehensive tumor response assessment but is too time-consuming for routine use.

Objective:
- Investigate whether RECIST long/short axes can serve as visual prompts for nnInteractive to enable fast and accurate tumor segmentation with improved inter-observer consistency.

Methods:
- Investigated on 5 datasets spanning CT (public) and MRI, including one multi-observer dataset
- RECIST long/short axes used as visual prompts combined with negative point prompts to guide nnInteractive segmentation.

Results:
- nnInteractive-based workflow significantly outperformed ellipsoid approximation in most metrics across all 5 datasets. It also achieved significantly higher inter-observer DSC and lower volume variability compared to clinical expert delineations, and comparable HD95.

Conclusion:
- RECIST-guided segmentation with nnInteractive enables accurate volumetric segmentation with minimal user input.
- Proposed workflow supports routine volumetric response assessment without added physician workload.

CONTACT

Denis Dudas
Dept. of Radiation Oncology, LMU University Hospital Munich
denis.dudas@med.uni-muenchen.de

INTRODUCTION

Background
- Tumor response in oncology is commonly evaluated using RECIST (Response Evaluation Criteria in Solid Tumors) guideline.
- RECIST assessment suffer from a considerable inter-observer variability (IOV) [1] and yield limited information regarding tumor morphology.
- Volumetric biomarkers could provide a more comprehensive assessment of tumor response.
- Unfortunately, full volumetric segmentation is not performed routinely as it is technically and time demanding.

Study Hypothesis
- RECIST long and short axis measurements (already available from routine response assessments) can serve as visual prompts for promptable segmentation models to automate volumetric GTV delineation.

nnInteractive Model for Segmentation
- nnInteractive – a promptable 3D segmentation model that accepts visual prompts [2].
- It generalizes across anatomical structures and imaging modalities without task-specific retraining.

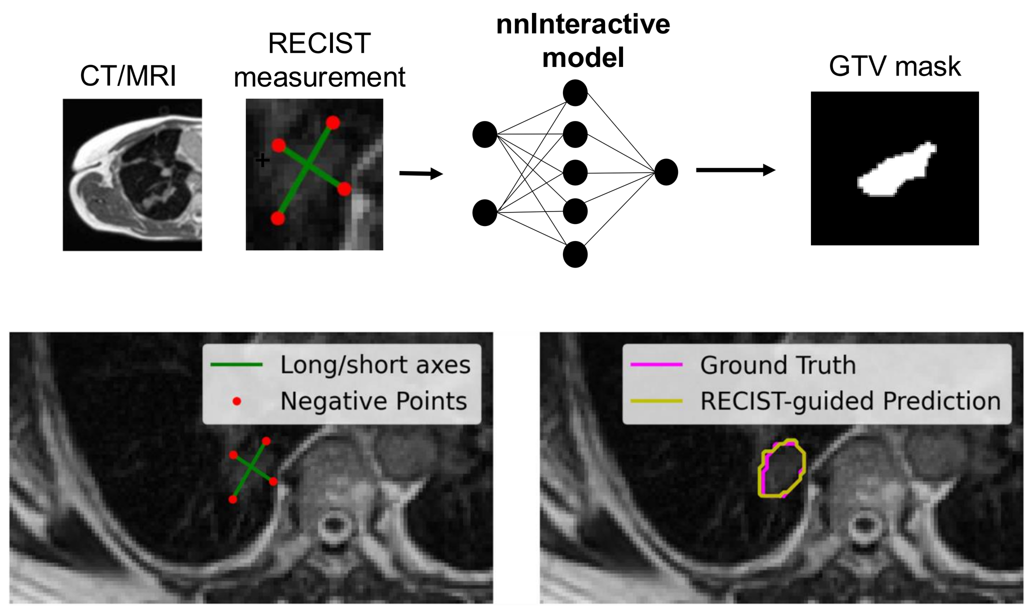

Figure 1: RECIST-guided GTV segmentation diagram (top row), and RECIST long/short axes measurements, including 4 negative points, used as prompts for nnInteractive model (bottom row).

METHODS AND MATERIALS

Prompting Strategy
- RECIST long/short axis measurements were derived from the ground truth GTV masks.
- Artificial IOV was introduced following values from the meta-analysis by Yoon et al. [1], and the resulting measurements were used as positive prompts for nnInteractive.
- To better constrain the boundary, negative point prompts were placed beyond the outer edge of each axis endpoint (see Figure 1).

Baseline comparison and evaluation
- Ellipsoid approximation, based on long/short axes measurements was used as a baseline to benchmark the proposed workflow.
- RECIST-guided predictions were compared with the ground truth mask or STAPLE consensus where multiple-observers were available
- utilizing Dice similarity coefficient (DSC), 95th percentile Hausdorff distance (HD95) and lesion volume.

Patient Cohorts
The proposed workflow was tested on 5 cohorts:

1. **Internal MRI dataset**: 116 samples (LMU)
2. **External MRI dataset**: 17 samples (AUMC)
3. **NSCLC-Radiomics CT dataset**: 422 samples (public [3])
4. **RIDER-LungCT-Seg (NSCLC-Seg) CT dataset**: 31 samples (public [4])
5. **NSCLC-Radiomics-Interobserver1 (NSCLC-Interobs) CT dataset**: 22 samples (public [4]), multiple observers

Results

- **Figure 2:** RECIST-guided GTV segmentation, nnInteractive vs ellipsoid approximation.
- **Table 1:** Statistical significance of differences between RECIST-guided nnInteractive masks and RECIST-guided ellipsoid approximation masks (Wilcoxon signed-rank test).
- **Figure 3**: IOV in NSCLC-Interobs dataset, RECIST-guided nnInteractive segmentation vs. manual clinical expert segmentation.

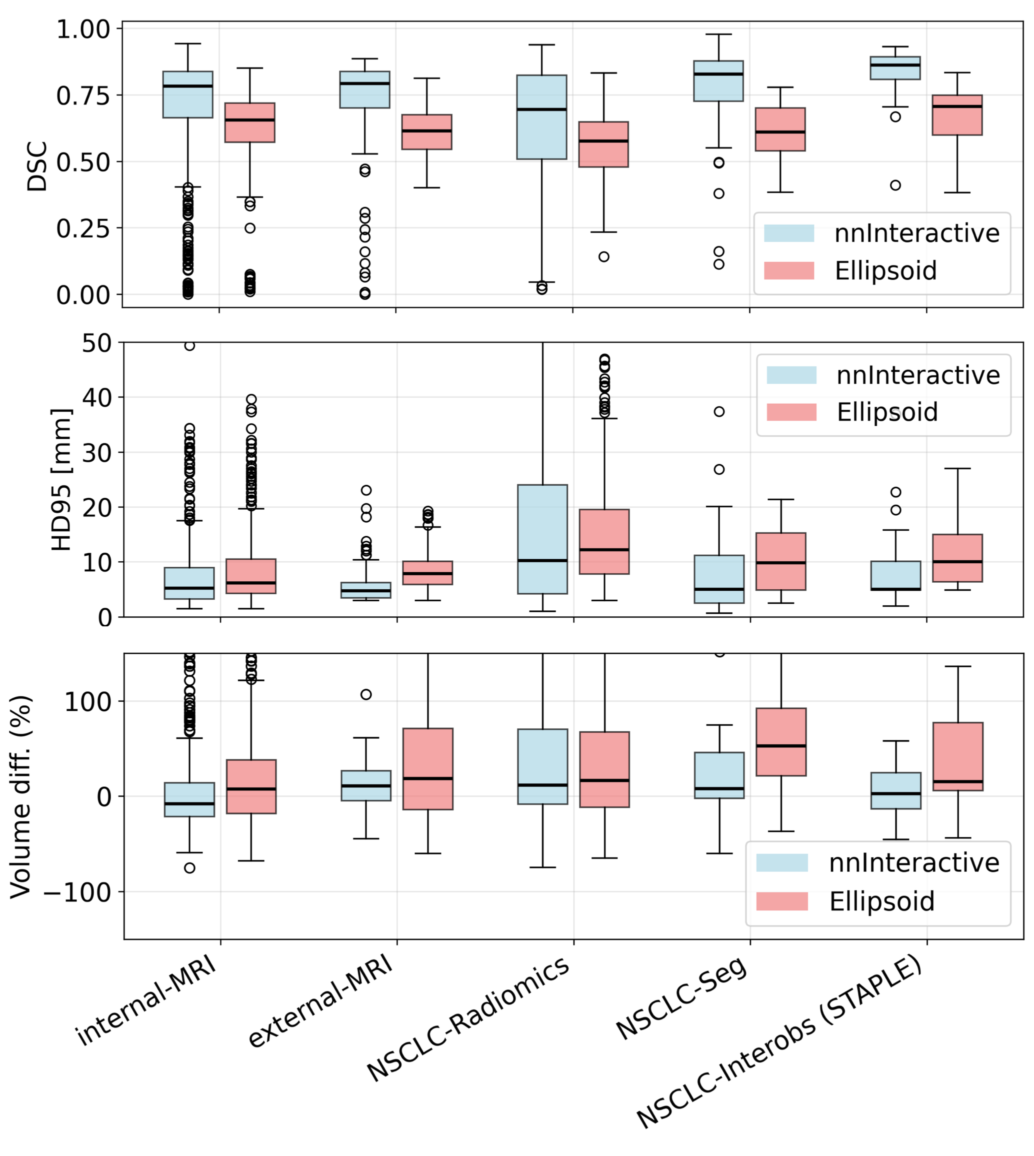

Figure 2: Comparison of DSC, HD95 and Volume diff. between the proposed RECIST-based nnInteractive workflow and RECIST-based ellipsoid approximation.

P values (Wilcoxon signed-rank test)

Dataset	DSC	HD95	Rel. Volume diff
Internal MRI	<0.001	<0.001	<0.001
External MRI	<0.001	<0.001	0.2
NSCLC-Radiomics	<0.001	0.6	0.21
NSCLC-Seg	<0.001	0.18	0.07
NSCLC-Interobs	<0.001	0.08	0.04

Table 1: p values (Wilcoxon signed-rank test, paired, two-sided) comparing RECIST-guided nnInteractive segmentation against ellipsoid approximation.

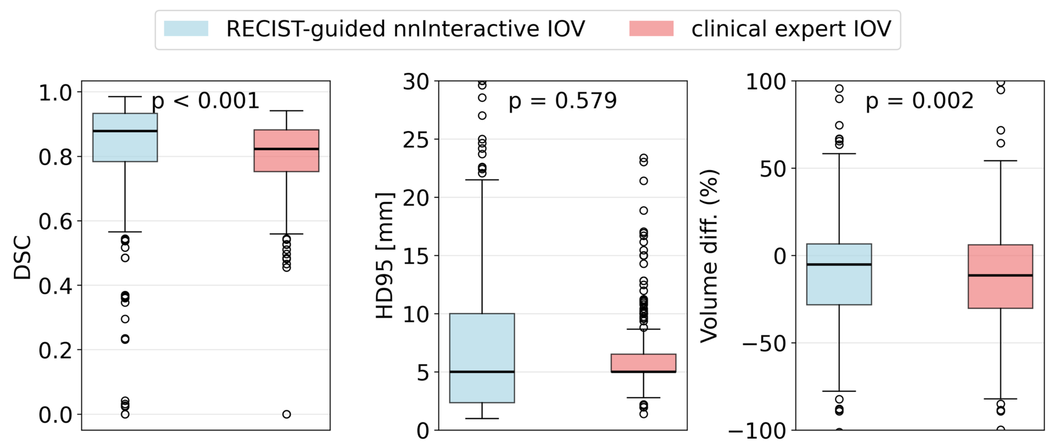

Figure 3: Pairwise inter-observer agreement in the NSCLC-Interobs dataset, including p values from Wilcoxon signed-rank test (paired, two-sided).

Conclusions

- RECIST measurements serve as effective visual prompts for nnInteractive-based tumor segmentation, with potential to improve routine volumetric response assessment.
- Strong segmentation performance across 5 independent CT & MRI datasets demonstrates robust cross-modality generalizability.
- RECIST-guided nnInteractive segmentation significantly outperforms ellipsoid approximation as a volumetric surrogate.
- RECIST-guided nnInteractive segmentation significantly improves inter-observer consistency compared to manual delineation.

REFERENCES

[1] Yoon SH et al. Observer variability in RECIST-based tumour burden measurements: a meta-analysis. Eur J Cancer 53 (2016). doi: 10.1016/j.ejca.2015.10.014

[2] Isensee F et al. nnInteractive: Redefining 3D Promptable Segmentation. arXiv (2025). doi: 10.48550/arXiv.2503.08373

[3] Aerts HJWL et al. Data From NSCLC-Radiomics (v4). The Cancer Imaging Archive. doi: 10.7937/K9/TCIA.2015.PF0M9REI

[4] Wee L et al. RIDER Lung CT Segmentation Labels from: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. The Cancer Imaging Archive. doi: 10.7937/tcia.2020.jit9grk8

[5] Wee L et al. Data from NSCLC-Radiomics-Interobserver1. The Cancer Imaging Archive. doi: 10.7937/tcia.2019.cwvlpd26