

Benefits of deep learning for delineation of organs at risk in head and neck cancer

van der Veen J.*¹, Willems S.*², Deschuymer S.¹, Robben D.^{2,3}, Crijns W.¹, Maes F.², Nuyts S.¹

¹KU Leuven, Dept. Oncology, Laboratory of Experimental Radiotherapy, & UZ Leuven, Radiation Oncology, B-3000 Leuven, Belgium ² KU Leuven, Dept. ESAT, Processing Speech and Images (PSI), & UZ Leuven, Medical Imaging Research Center, B-3000 Leuven, Belgium ³ Icometrix, B-3000 Leuven, Belgium

* These authors contributed equally to this work

BACKGROUND & AIM

Precise delineation of organs at risk (OARs) in head and neck cancer (HNC) is necessary for accurate radiotherapy. Although guidelines exist, significant interobserver variability (IOV) remains.

Aim: Create and validate a 3D convolutional neural network for semi-automated delineation of OARs

1. time efficiency
2. network accuracy
3. IOV compared to manual delineation

CONCLUSION

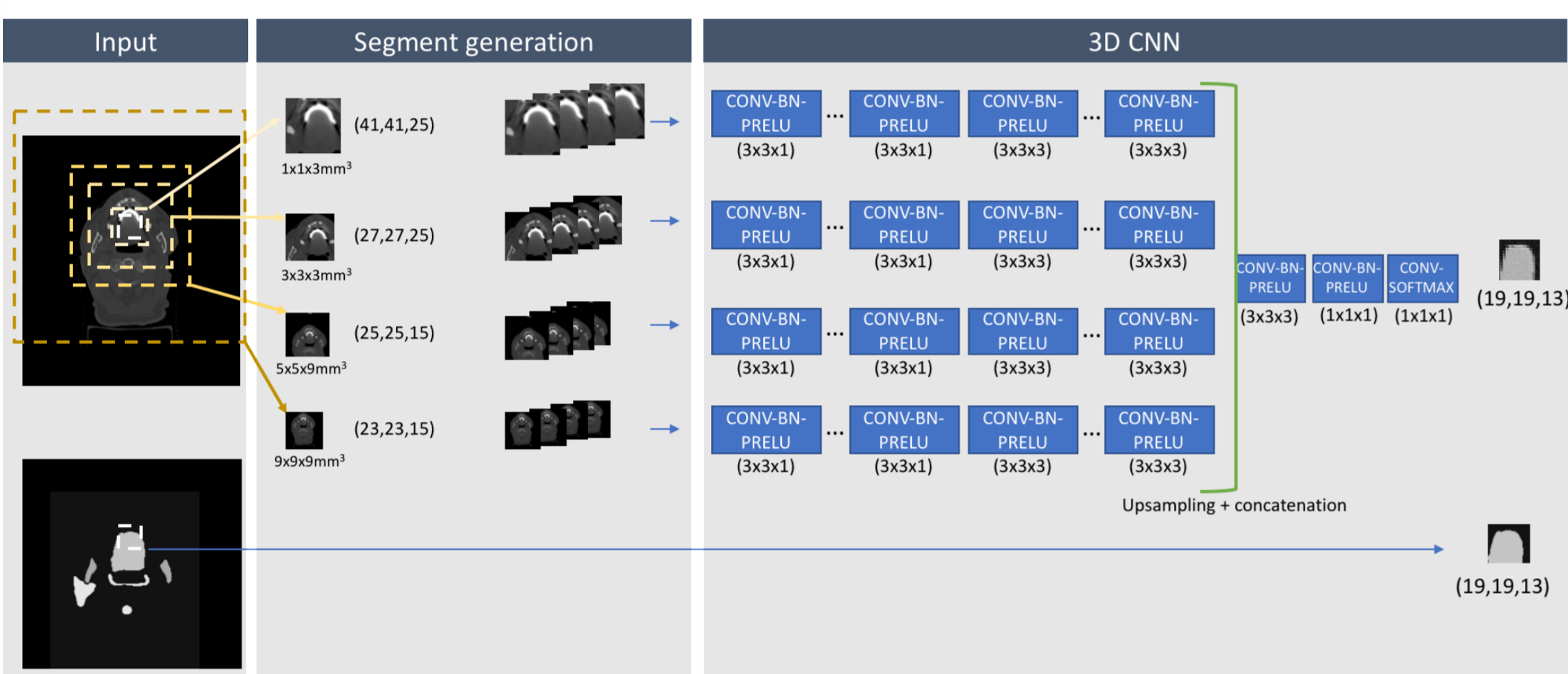
We created a CNN for semi-automated delineation of the most important OARs in HNC patients, based on international consensus guidelines. Two RO **validated** the CNN in 15 new HNC patients.

Semi-automated delineation using the CNN compared to manual delineation

1. **is faster**
2. **creates OARs accurately.** Only minor corrections needed
3. **reduces IOV**

MATERIAL AND METHODS

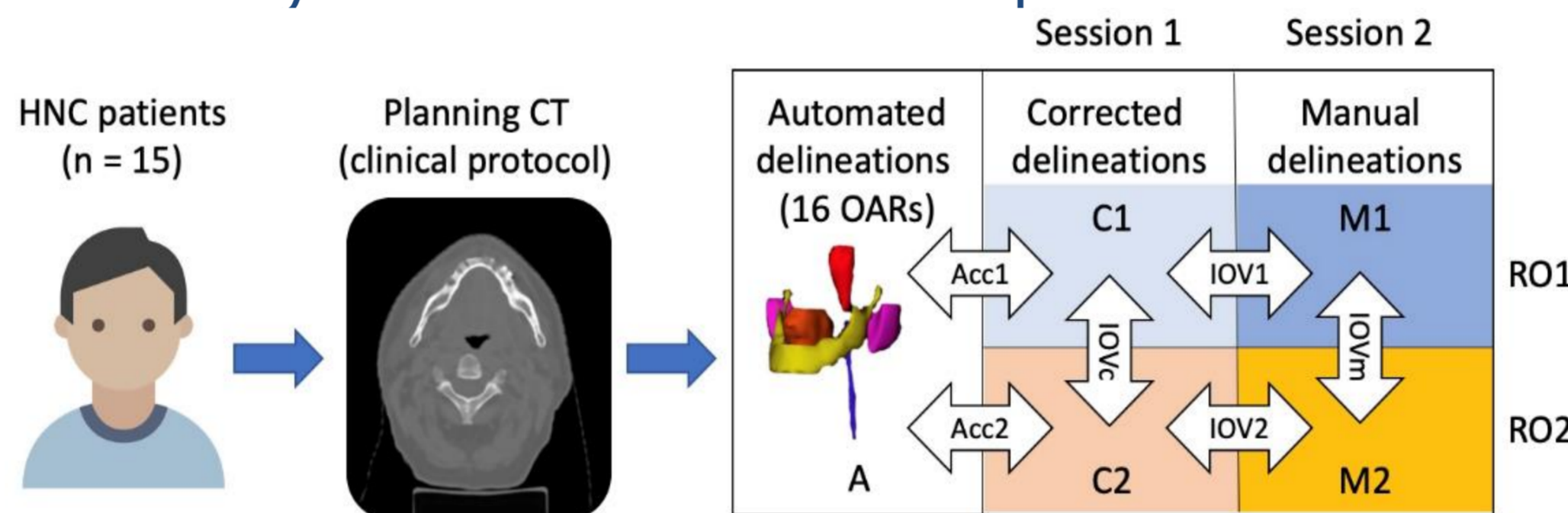
3D convolutional neural network (CNN)



The CNN was made using delineations of OARs in 70 HNC patients, delineated using guidelines from Brouwer et al. (2015)^α and Christianen et al. (2011)^β.

Study design

Automated delineations of 16 OARs in 15 new HNC patients. 2 radiation oncologists (RO1, RO2) corrected automated delineations and manually delineated the same patients.



Acc1, Acc2: Accuracy of automated delineations; IOV: Inter-observer variability between corrected delineations; IOVm: Inter-observer variability between manual delineations; IOV1, IOV2: Intra-observer variability between corrected and manual delineations

Delineation agreement analysis

Dice similarity coefficient (DSC): measure for overlap between two delineations A and B

$$DSC = 2 \times \frac{|A \cap B|}{|A| + |B|} \times 100\% \quad (\text{ideal value} = 100\%)$$

Average symmetric surface distance (ASSD): mean distance between delineations A and B

$$ASSD(A, B) = \frac{h(A, B) + h(B, A)}{2} \quad (\text{ideal value} = 0 \text{ mm})$$

Used to calculate **network accuracy** (difference between automated and corrected delineations (Acc) and **IOV** between both ROs (IOVc and IOVm).

RESULTS

1. Time efficiency

Correction time of automated delineations faster than manual delineations

- RO1: 17 vs 30 minutes
 - RO2: 27 vs 38 minutes
- Mean of RO1, RO2: 23 vs 34 minutes, **p<10-6** Time gain: 6 - 19 minutes (33% on average)

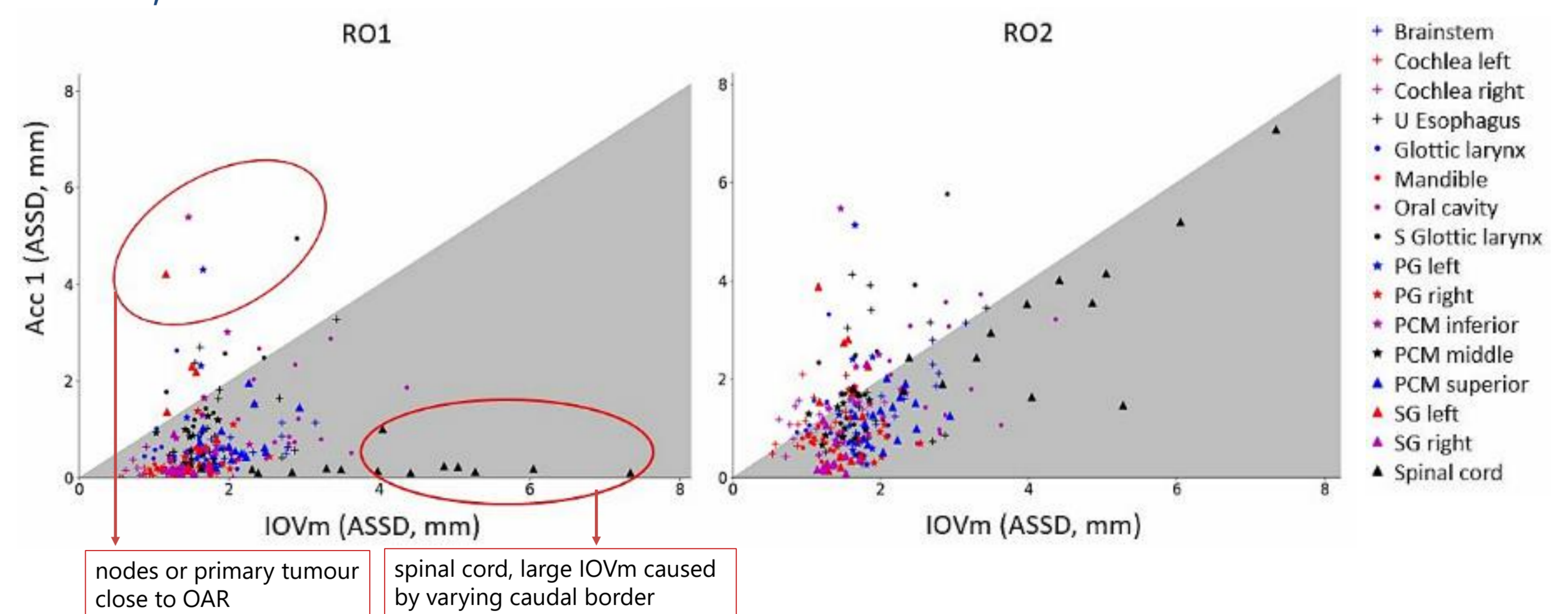
2. Network Accuracy

Automated vs corrected delineation for each RO separately (Acc1, Acc2). All values are reported as mean ± standard deviation for all patients (n=15) and for both observers for volumes.

	Acc1 (RO1)		Acc2 (RO2)		Volume (ml)		
	DSC (%)	ASSD (mm)	DSC (%)	ASSD (mm)	Manual	Automated	Corrected
Brainstem	94.9±2.1	0.7±0.2	98.1±1.6	1.7±0.6	228.8±44.4	263.1±32.4	231.1±41.7
Cochlea left	98.7±1.1	0.1±0.1	95.1±7.2	1.1±0.5	0.8±0.2	1.6±0.6	1.2±0.4
Cochlea right	96.4±8.4	0.2±0.2	80.0±25.8	1.1±0.4	1.0±0.3	1.4±0.7	1.2±0.4
Upper Esophagus	83.1±17.2	1.3±0.9	79.3±17.9	2.0±1.3	49.4±17.8	49.8±17.5	49.4±14.7
Glottic larynx	76.8±16.1	0.7±0.6	64.4±14.7	1.3±0.6	38.0±15.9	26.6±11.4	35.8±15.7
Mandible	98.8±0.6	0.2±0.1	91.2±2.2	0.7±0.2	620.2±106.2	596.5±116.7	619.3±117.2
Oral cavity	93.1±7.3	1.5±0.8	85.4±10.1	2.1±1.0	1040.0±196.3	1025.3±234.0	1087.8±223.6
Supraglottic larynx	74.2±23.7	1.5±1.1	67.1±23.5	2.1±1.3	160.6±49.5	124.2±47.9	156.8±45.0
Parotid gland left	96.1±2.9	0.9±1.1	93.4±5.3	1.3±1.3	276.9±86.1	275.5±81.1	276.5±86.5
Parotid gland right	96.1±3.3	0.5±0.3	92.9±5.2	0.9±0.5	303.1±104.0	281.7±88.4	290.3±95.7
PCM inferior	80.1±23.7	1.3±1.3	71.3±21.0	1.7±1.1	43.8±9.7	41.7±15.9	44.4±13.3
PCM middle	84.1±1.0	0.6±0.3	76.4±10.0	1.1±0.4	50.3±17.9	43.2±13.8	48.2±16.7
PCM superior	82.9±15.4	0.8±0.5	75.5±13.0	1.2±0.4	84.4±25.2	73.5±26.4	80.3±21.0
Submandibular gland left	96.5±4.9	1.0±1.2	91.4±8.3	1.4±1.2	91.1±28.5	90.3±41.0	87.7±27.9
Submandibular gland right	98.2±2.7	0.3±0.2	95.1±7.0	0.6±0.6	92.7±31.5	77.2±31.3	85.4±28.4
Spinal cord	97.8±3.1	0.2±0.2	92.8±3.6	3.6±2.3	136.3±23.7	166.8±32.8	152.2±28.5

PCM: pharyngeal constrictor muscles; mm: millimetre; ml: millilitre

Corrections needed by RO1 and RO2 vs manual IOV. Each data point represents an OAR from one patient. For all structures in the grey zone, the corrections are smaller than IOVm.



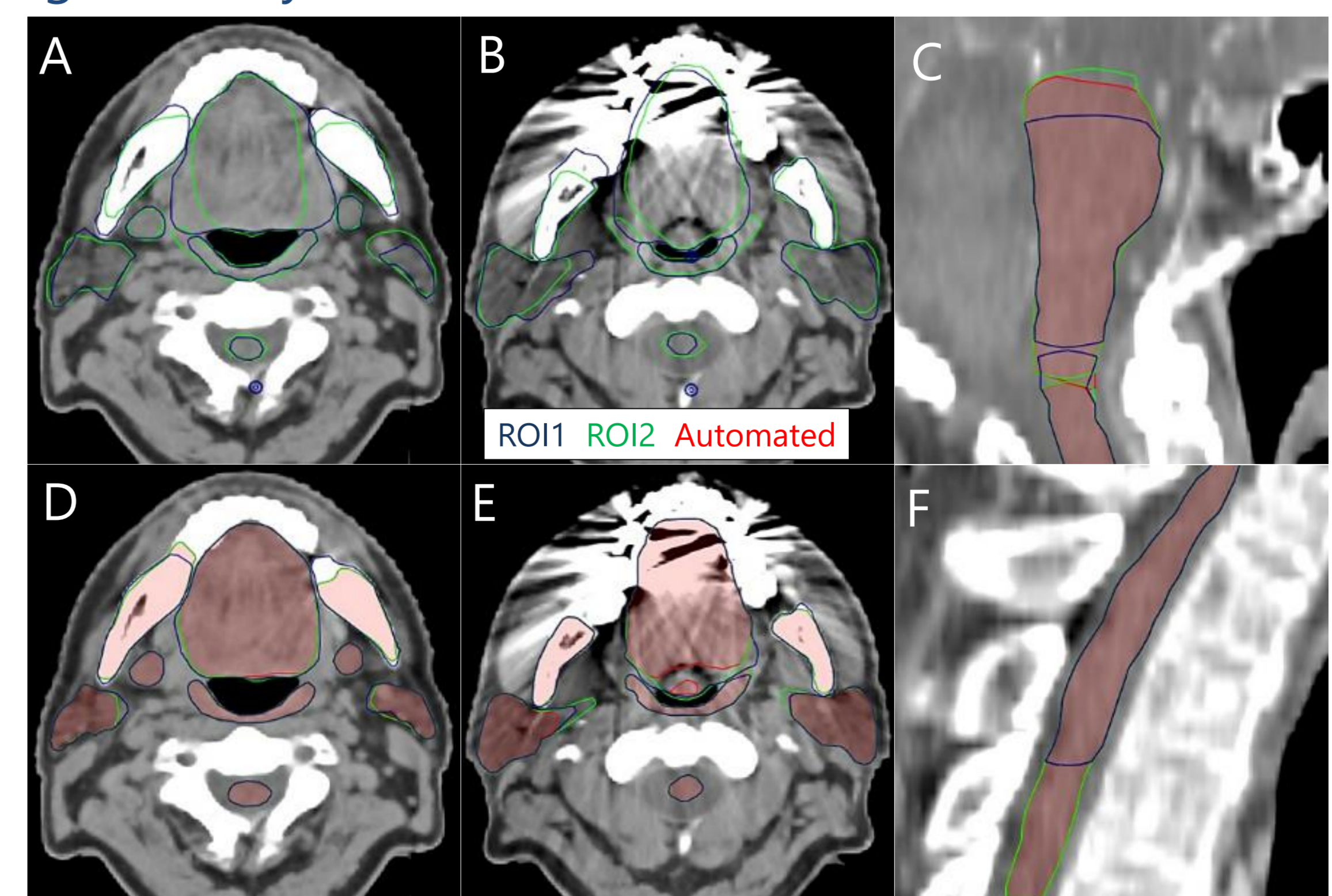
3. Interobserver variability

RO1 compared to RO2 for each OAR for manual and corrected delineations separately. All values are reported as mean ± standard deviation for all patients (n=15). Statistically significant improvement (p<0.05) in IOV for corrected vs manual delineations (IOVc vs IOVm) indicated in bold, in favour of the best method.

	Interobserver variability (RO1 vs RO2)			
	IOVm (manual)		IOVc (corrected)	
	DSC (%)	ASSD (mm)	DSC (%)	ASSD (mm)
Brainstem	68.1±11.7	2.2±0.5	70.8±11.9	1.6±0.6
Cochlea left	48.7±13.4	1.1±0.3	53.9±17.7	1.1±0.5
Cochlea right	51.4±12.0	1.1±0.4	63.3±19.8	1.0±0.4
Upper Esophagus	64.3±15.3	2.0±0.6	79.3±16.5	1.1±0.7
Glottic larynx	73.1±17.8	1.4±0.4	89.8±9.0	0.9±0.3
Mandible	93.7±2.9	1.2±0.2	98.8±1.4	0.7±0.2
Oral cavity	94.0±4.7	2.9±0.6	96.2±3.2	1.6±0.7
Supraglottic larynx	86.3±10.9	1.8±0.4	93.9±3.2	1.3±0.4
Parotid gland left	88.8±3.4	1.7±0.1	93.7±5.2	0.8±0.6
Parotid gland right	89.1±3.3	1.8±0.3	95.1±3.8	0.7±0.4
PCM inferior	76.0±8.2	1.7±0.3	87.7±7.8	1.2±0.5
PCM middle	71.5±7.2	1.5±0.2	86.4±8.2	0.8±0.3
PCM superior	53.5±8.0	2.1±0.3	77.6±12.7	1.2±0.5
Submandibular gland left	86.5±6.5	1.5±0.2	92.7±7.9	0.8±0.6
Submandibular gland right	88.4±3.9	1.4±0.2	96.0±5.3	0.4±0.6
Spinal cord	70.8±6.8	4.4±1.9	75.7±6.5	3.7±2.3

PCM: pharyngeal constrictor muscles; mm: millimetre

Decrease in IOV (A and B vs D and E). RO1 and RO2 used different cranial and caudal borders for brainstem and spinal cord (C and F), explaining relatively low DSC in IOV.



Corresponding author: julie.vanderveen@uzleuven.be & siri.willems@kuleuven.be

References: ^α CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. Brouwer, C. et al. Radiother Oncol, 2015; 117: 83-90. ^β Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. Christianen M. et al. Radiother Oncol. 2011; 101: 394-402