PARROT: A Web Platform for AI-Driven Radiation Therapy Planning

Margerie Huet Dastarac[1], Wei Zhao[2], Sylvain Deffet[3], Eléonore Longton[4], Sophie Cvilic[4], Edmond Sterpin[1], John Lee[1], and Ana Barragan Montero[1]

[1]Molecular Imaging, Radiation and Oncology (MIRO) Laboratory, Institut de Recherche Expérimentale et Clinique (IREC), UCLouvain, Belgium
[2]Institute of Information and Communication Technologies (ICTEAM), UCLouvain, Belgium
[3]Ion Beam Applications SA, Louvain-La-Neuve, Belgium
[4]Department of Radiation Oncology, Saint-Luc University Hospital, Belgium

Abstract Artificial Intelligence (AI) is gaining momentum in medical fields like radiation therapy. Models to delineate structures and predict optimal dose trade-off for patients are developed by the research community. Although some of these models are starting to be implemented clinically, an end-to-end workflow with a user-friendly graphical interface to visualize the result of successive AI models and to support decision of treatment indication is still prospective.

To address this problem, we present PARROT, a free and open-source web platform that facilitates the use of AI delineation and dose prediction models and the visualization of the models outputs. The treatment decision support shows clinical evaluation tools to compare dose distributions, among which normal tissue complication probabilities (NTCP) models.

1 Introduction

Artificial Intelligence (AI) is gaining momentum in medical fields like radiation therapy. Models to delineate structures and predict optimal dose trade-off for patients are developed by the research community. Although some of these models are starting to be implemented clinically, an end-to-end workflow with a user-friendly graphical interface to visualize the result of successive AI models and to support decision of final treatment indication is still prospective.

To address this problem, we introduce PARROT which stands for Platform for ARtificial intelligence guided Radiation Oncology Treatment. PARROT is a user-friendly, free and open-source web platform which allows users to visualize DICOM files, run AI models in a single mouse click, display and evaluate predictions easily. Several trained state-of-the-art dose prediction and contour segmentation models are available. Moreover, users can add their own models from the embedded code editor. The app consists of a front-end built in React JavaScript including the graphical interface with DICOM viewer, a selection panel of AI models, an editor of contours predictions and evaluation tools. The back-end is built with Flask to handle the predictions of AI models.

Download instructions are available on our website https://huggingface.co/spaces/AI4MIRO/parrot.ai and the source code can be found on the Gitlab repository https://gitlab.com/ai4miro/parrot.

2 Materials and Methods

The interface is organized in four screens, Figure 1.

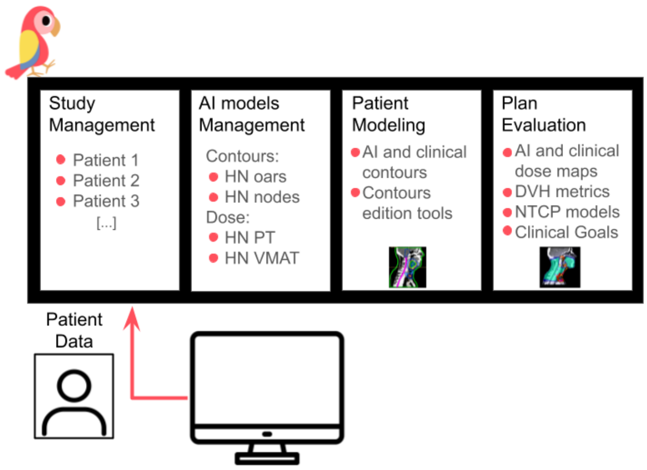


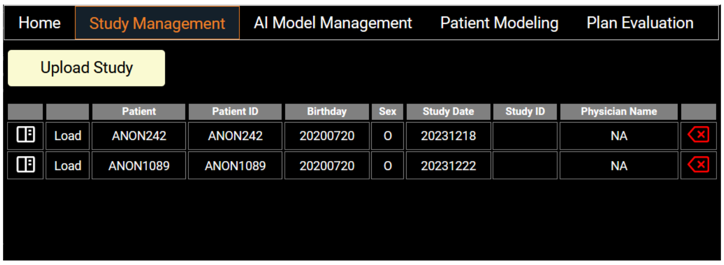
Figure 1: Overview of the PARROT platform.

2.1 Study Management

Users can import patient data from this screen to a Orthanc daemon—a lightweight, standalone DICOM server [1]. It runs locally and therefore ensures the confidentiality of patient files. The interactions with the Orthanc daemon are done through its RESTful API [2]. The list of patient data present in the Orthanc server is displayed with summarized information (Figure 2). Users can load the data into the app by clicking on the Load button. The left icon displays more complete information from DICOM tags (Figure 3) and users can delete a patient study from the server with the red arrow on the right.

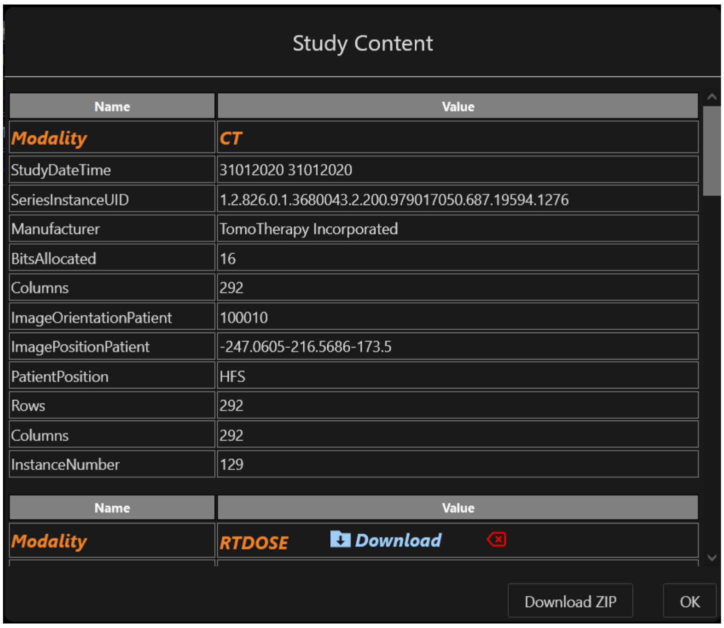
2.2 AI Models Management

Eight segmentation models and six dose prediction models are made available and listed on the screen (Figure 4). These models are trained from state-of-the-art architectures such as nnUNet [3], SwinUNETR [4] and HDUNet [5] (Table 1). On the left hand side of each line, a standardized description of the model is available. The PDF description file contains details about the architecture with a link of the original research



		Patient	Patient ID	Birthday	Sex	Study Date	Study ID	Physician Name	
	Load	ANON242	ANON242	20200720	O	20231218		NA	
	Load	ANON1089	ANON1089	20200720	O	20231222		NA	

Figure 2: Study management screen.

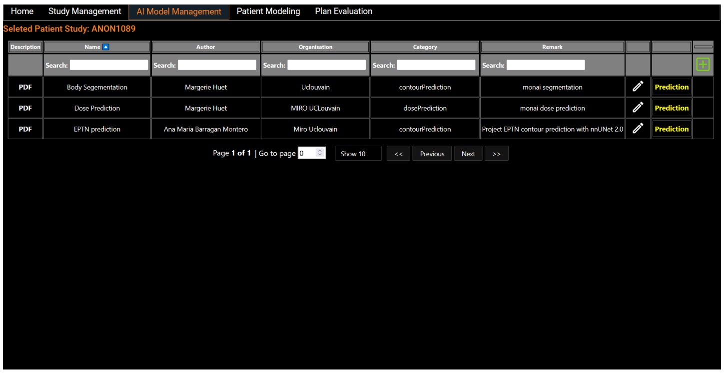

Study Content

Name	Value
Modality	*CT*
StudyDateTime	31012020 31012020
SeriesInstanceUID	1.2.826.0.1.3680043.2.200.979017050.687.19594.1276
Manufacturer	TomoTherapy Incorporated
BitsAllocated	16
Columns	292
ImageOrientationPatient	100010
ImagePositionPatient	-247.0605-216.5686-173.5
PatientPosition	HFS
Rows	292
Columns	292
InstanceNumber	129

Name	Value
Modality	*RTDOSE* *Download*

Figure 3: Content information of the study stored in DICOM tags.

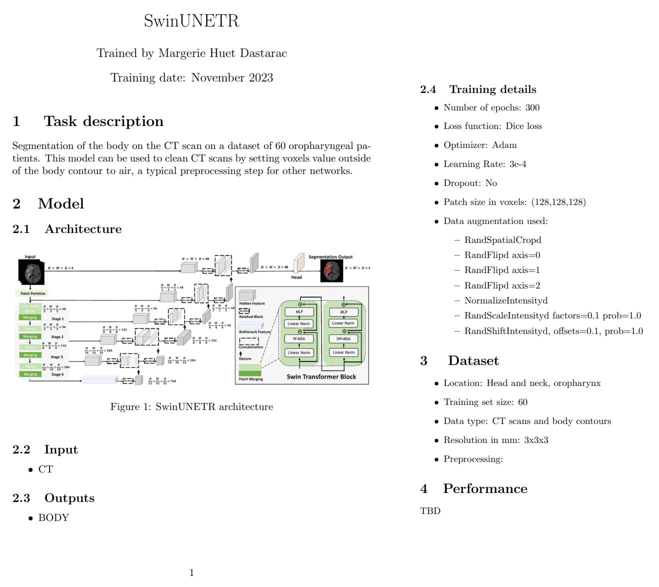
article, details on the training process including dataset, and performance on our test set (Figure 5).



Seleted Patient Study: ANON1089

Description	Name	Author	Organisation	Category	Remark		
	Search:	Search:	Search:	Search:	Search:		
PDF	Body Segementation	Margerie Huet	Uclouvain	contourPrediction	monai segmentation		Prediction
PDF	Dose Prediction	Margerie Huet	MIRO UCLouvain	dosePrediction	monai dose prediction		Prediction
PDF	EPTN prediction	Ana Maria Barragan Montero	Miro Uclouvain	contourPrediction	Project EPTN contour prediction with nnUNet 2.0		Prediction

Figure 4: AI Models Management screen. A list of the embedded AI models are displayed.


SwinUNETR

Trained by Margerie Huet Dastarac

Training date: November 2023

1 Task description

Segmentation of the body on the CT scan on a dataset of 60 oropharyngeal patients. This model can be used to clean CT scans by setting voxels value outside of the body contour to air, a typical preprocessing step for other networks.

2 Model

2.1 Architecture

Figure 1: SwinUNETR architecture

2.2 Input

- CT

2.3 Outputs

- BODY

2.4 Training details

- Number of epochs: 300
- Loss function: Dice loss
- Optimizer: Adam
- Learning Rate: 3e-4
- Dropout: No
- Patch size in voxels: (128,128,128)
- Data augmentation used:
  - RandSpatialCropd
  - RandFlipd axis=0
  - RandFlipd axis=1
  - RandFlipd axis=2
  - NormalizeIntensityd
  - RandScaleIntensityd factors=0.1 prob=1.0
  - RandShiftIntensityd, offsets=0.1, prob=1.0

3 Dataset

- Location: Head and neck, oropharynx
- Training set size: 60
- Data type: CT scans and body contours
- Resolution in mm: 3x3x3
- Preprocessing:

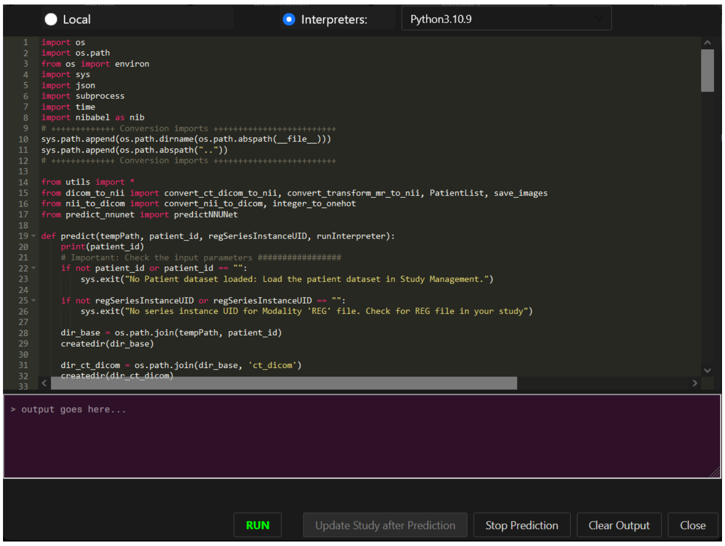
4 Performance

TBD

1

Figure 5: Example of PDF with description of a specific model provided in PARROT.

To run a model prediction, the user simply clicks on the yellow Prediction button on the right hand side of the model line. This will open a Python scripting editor to run the code (Figure 6). Users can select a Python environment with a set of libraries from the drop-down menu in the upper right corner. By hovering on an item, a description of the libraries it contain is displayed. Once the prediction is done, users can upload the result to the study to visualize the output in the Patient Modeling screen (Figure 8). Users also have the flexibility to run their own models by loading their processing and inference scripts in the Python scripting editor. On our website, we provide a description and examples using the API to retrieve the patient information loaded on PARROT and upload the result.

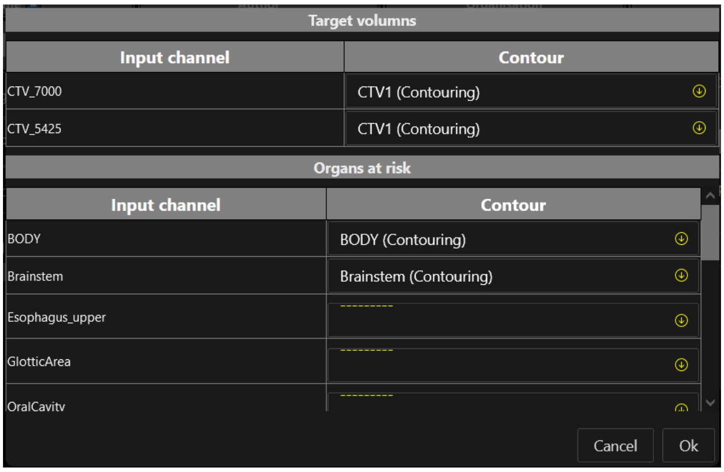


```python
import os
import os.path
from os import environ
import sys
import json
import subprocess
import time
import nibabel as nib
# +++++++++++++ Conversion imports ++++++++++++++++++++++++
sys.path.append(os.path.dirname(os.path.abspath(__file__)))
sys.path.append(os.path.abspath(".."))
# +++++++++++++ Conversion imports ++++++++++++++++++++++++

from utils import *
from dicom_to_nii import convert_ct_dicom_to_nii, convert_transform_mr_to_nii, PatientList, save_images
from nii_to_dicom import convert_nii_to_dicom, integer_to_onehot
from predict_nnunet import predictNNUNet

def predict(tempPath, patient_id, regSeriesInstanceUID, runInterpreter):
    print(patient_id)
    # Important: Check the input parameters #################
    if not patient_id or patient_id == "":
        sys.exit("No Patient dataset loaded: Load the patient dataset in Study Management.")

    if not regSeriesInstanceUID or regSeriesInstanceUID == "":
        sys.exit("No series instance UID for Modality 'REG' file. Check for REG file in your study")

    dir_base = os.path.join(tempPath, patient_id)
    createdir(dir_base)

    dir_ct_dicom = os.path.join(dir_base, 'ct_dicom')
    createdir(dir_ct_dicom)
```

```
> output goes here...
```

Figure 6: Python scripting editor. Users can edit their script, run and display output from this window.

We distinguish two categories of models: the segmentation models and the dose prediction models. A key difference between the two categories is that segmentation models require only the anatomy of the patient (CT scan or MRI image) whereas dose prediction models also require binary masks of tumor volumes (TV) and organs at risk (OARs). For the latter, we implemented a dialog that allows users to map structures present in the current patient data and the input required for the AI dose prediction model, Figure 7.



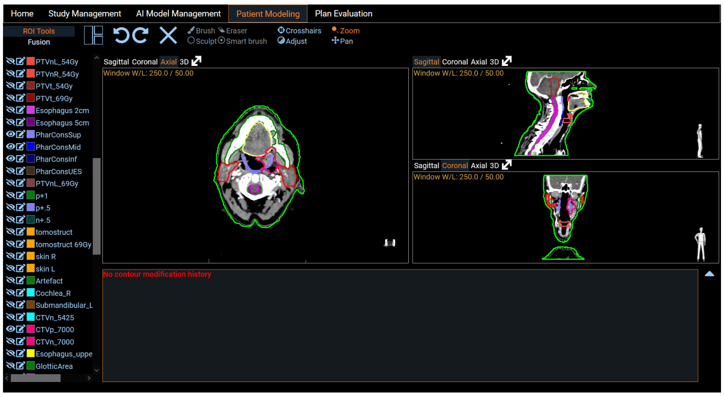
Target volumns	
Input channel	**Contour**
CTV_7000	CTV1 (Contouring)
CTV_5425	CTV1 (Contouring)

Organs at risk	
Input channel	**Contour**
BODY	BODY (Contouring)
Brainstem	Brainstem (Contouring)
Esophagus_upper	---------
GlotticArea	---------
OralCavity	---------

Figure 7: Mapping dialog between structures present in the current patient data and required inputs of the model. The left column displays the input channels required for that specific AI model and the right column a drop-down list of the contours available in the loaded patient study.

2.3 Patient Modeling

This screen allows an interactive visualization of contours, whether uploaded by the user or generated by AI models in the previous tab. Similarly to the visualization of treatment planning software, users can scroll through the anatomy and contours of the patient, zoom, pan and change window range and levels. Users can edit contours with the different tools that can be seen Figure 8. The platform maintains a detailed history of modifications, and users can provide descriptive labels for refining AI models.



No contour modification history

Figure 8: Patient Modeling screen.

2.4 Plan Evaluation

A comprehensive display of two dose distributions comparison is presented on the last screen. The two dose distributions can originate from the importation step or result from the inference of AI models. We provide a set of tools for comparison, such as dose-volume histograms (DVH) curves, dose statistics, and the fulfillment of clinical goals. Clinical goals can be uploaded following an Excel template so that it fits any clinical requirements of the users institutions. The platform incorporates treatment indications based on NTCP protocols. It has been shown in multiple studies that predicted dose distributions yield meaningful output when used with NTCP models for treatment selection decision support [6, 7]. Two NTCP protocols, validated by the clinical community have been integrated, one for head and neck cancer patients (with four NTCP models for xerostomia grade $\geq$ II and grade $\geq$ III and dysphagia grade $\geq$ II and grade $\geq$ III)[8] and one for esophagus cancer patients to estimate pulmonary complications probabilities [9]. This is a useful decision support tool for the clinical community.

Priority	ROI/POI	Clinical Goals	Dose 1	Dose 2
4-Less Important	Mandible	Mandible D2%<7000cGy	54.28Gy	48.71Gy
3-Important	Esophagus	Esophagus Dmean<4500cGy	26.19Gy	28.41Gy
3-Important	Esophagus	Esophagus D30%<6000cGy	40.09Gy	38.72Gy
3-Important	Esophagus	Esophagus D0,03cc<6900cGy	55.63Gy	46.59Gy
1-Most Important	SpinalCord	SpinalCord D0,03cc<4500cGy	13.58Gy	19.65Gy
3-Important	Parotid_R	Parotid_R Dmean<2000cGy	30.70Gy	27.95Gy
3-Important	Parotid_L	Parotid_L Dmean<2000cGy	29.36Gy	30.16Gy
4-Less Important	Submandibular_L	Submandibular_L Dmean<3500cGy	46.90Gy	37.44Gy
4-Less Important	Submandibular_R	Submandibular_R Dmean<3500cGy	62.74Gy	50.45Gy
4-Less Important	OralCavity	OralCavity Dmean<4000cGy	18.97Gy	29.99Gy
3-Important	SpinalCord_PRV	SpinalCord_PRV D0,03cc<4500cGy	19.97Gy	23.32Gy

Model Name: ntcpLIPPv22

NTCP model criteria:	Value:
ΔNTCP Xerostomia grade >= II >= 10%	5.32%
ΔNTCP Xerostomia grade >= III >= 5%	1.89%
ΔNTCP Dysphagia grade >= II >= 10%	-13.65%
ΔNTCP Dysphagia grade >= III >= 5%	-9.63%
SumΔNTCP grade >= II >= 15% and each >= 5%	False
SumΔNTCP grade >= III >= 7.5% and each >= 2.5%	False

This patient is selected for a XT treatment according to Landelijk Indication Protocol Protontherapy (LIPPv2.2)

Figure 9: Plan Evaluation screen.

3 Results

Three main use cases that can benefit from PARROT are the following.

3.1 Visualization of state-of-the-art AI models on patient of local institute

The integration of a novel AI model within a clinical setting necessitates a careful and meticulous analysis of the tool. It is imperative to conduct thorough testing using a local patient database to compare the performance of the model in its initial testing conditions and the specific context of the new healthcare institution. The PARROT framework facilitates seamless visualization of predictions, eliminating the need for manual coding.

3.2 Curate own AI segmentation models with the contour editing tools

Utilizing the features available on the viewer screen, clinicians have the capability to rectify contours produced by AI models and annotate the rationale behind the modifications. This process helps to identify systematic errors and limitations of the AI models, guiding efforts to curate the training set and retrain the model.

Segmentation		Dose prediction	
Location	Architecture	Location	Architecture
Head and Neck Nodes	nnUNetv2	Head and Neck PT	HDUNet
Head and Neck OARs	nnUNetv2	Head and Neck VMAT	HDUNet
Prostate OARs	nnUNetv2	Prostate IMRT	HDUNet
Breast OARs	nnUNetv2	Breast PT	HDUNet
Brain OARs	nnUNetv2	Esophagus IMRT	HDUNet
Brain Tumor	SwinUNETR	Esophagus PT	HDUNet
Esophagus OARs	nnUNetv2		
Body	SwinUNETR		

Table 1: List of provided AI models.

3.3 Decision support between PT and XT treatments

The dose evaluation screen of PARROT serves as a decision support tool for choosing between PT and XT treatments. Users can have a seamless overview of DVH curves, dose metrics, clinical goals reached and indication of NTCP models. In the current version, AI dose prediction models and corresponding NTCP models are provided for oropharyngeal cancer patients (PT versus VMAT) and esophageal cancer (PT versus IMRT).

4 Discussion

We believe that PARROT can have a positive impact on the safe integration of AI models in the clinic. On the one hand, PARROT can facilitate model comparison, sharing, and benchmarking by the research community. On the other hand, it can foster the dissemination of AI models to the clinics, standardizing and improving the medical practice. In addition, it can be a valuable tool to foster consensus between experts when updating or creating new contouring and planning guidelines. In the future, we would like to extend PARROT to integrate different methods of uncertainty quantification both for segmentation and dose prediction models. These uncertainty estimation tools can be used to curate AI models and datasets and flag out-of-distribution patients and therefore indicate clinicians that the AI model should not be used for that specific patient.

Currently, the platform supports CT and MR images, and automatically applies a registration field whenever a REG DICOM file is uploaded. But future releases will allow PET scan to be viewed and used by AI models.

5 Conclusion

PARROT is a a free and open-source web platform that facilitates the use of AI models in the radiation therapy field. The web platform allows users to visualize DICOM files, run AI models (contour segmentation and dose prediction models) in a single mouse click, display and evaluate predictions easily. PARROT offers clinical evaluation tools to compare dose distributions and support treatment decision. Among these tools, users can find dose-volume histograms curves, dose statistics, validation of clinical goals and outputs of normal tissue complication probabilities models. We believe that PARROT can have a positive impact on the safe integration of AI models in the clinic.

References

[1] S. Jodogne. “The Orthanc Ecosystem for Medical Imaging”. *JDI* 31.3 (May 2018).

[2] *REST API of Orthanc — Orthanc Book documentation*. `https://orthanc.uclouvain.be/book/users/rest.html`. Accessed: 2024-1-8.

[3] F. Isensee, P. F. Jaeger, S. A. A. Kohl, et al. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation”. *Nat. Methods* 18.2 (Feb. 2021).

[4] A. Hatamizadeh, V. Nath, Y. Tang, et al. “Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images”. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Springer International Publishing, 2022.

[5] D. Nguyen, X. Jia, D. Sher, et al. “3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture”. *PMB* 64.6 (Mar. 2019).

[6] M. Huet-Dastarac, S. Michiels, S. T. Rivas, et al. “Patient selection for proton therapy using Normal Tissue Complication Probability with deep learning dose prediction for oropharyngeal cancer”. *Med. Phys.* (May 2023).

[7] C. Draguet, A. M. Barragán-Montero, M. C. Vera, et al. “Automated clinical decision support system with deep learning dose prediction and NTCP models to evaluate treatment complications in patients with esophageal cancer”. *Radiother. Oncol.* 176 (Nov. 2022).

[8] N. V. voor Radiotherapie en Oncologie (NVRO). *LIPPv2*. `https://nvro.nl/images/documenten/rapporten/2019-08-15__Landelijk_Indicatieprotocol_Protonentherapie_Hoofdhals_v2.2.pdf`. [Accessed 25-01-2024].

[9] M. Thomas, G. Defraene, M. Lambrecht, et al. “NTCP model for postoperative complications and one-year mortality after trimodality treatment in oesophageal cancer”. *Radiother. Oncol.* 141 (Dec. 2019).