Appendix A. Network architecture

The general network architecture was adapted from Isola et al. (2017). Contoured CT slices were used as input to the generator as 3-channel, 128×128 images. We used a U-net architecture, where the generator was comprised of an encoder and a decoder stage. We used 4×4 2D convolutions with stride 2 and padding 1. Each convolution layer was followed by a leaky ReLU and batch normalization. Deconvolution layers were followed by 50% dropout, ReLU, and batch normalization.

The encoder consisted of four downsampling layers. The first generated 64 channels, and each subsequent layer downsampled by a factor of 2. This was followed by 2 bottleneck layers, before the data was then passed through 4 upsampling layers. The output of each downsample layer was concatenated to the input of the corresponding upsample layer. The final output was a 3-channel, 128×128 slice.

The decoder consisted of five convolution layers, where the first four each downsample the output by 2. The fifth, and last layer, mapped to a scalar output. Once again, we applied batch normalization and leaky ReLU after the first four layers. The final layer was passed through sigmoid activation.

Appendix B: Random forest architecture

Feature	Description
Structure	Structure that the voxel is classified as
y-coordinate	Voxel's positions on the y-axis in a slice
z-coordinate	Plane of voxel's slice
Distance to larynx	Shortest path between voxel and the surface of the larynx
Distance to esophagus	Shortest path between voxel and the surface of the esophagus
Distance to limPostNeck	Shortest path between voxel the surface of the limPostNeck
Distance to PTV56	Shortest path between voxel and the surface of the PTV56
Distance to PTV63	Shortest path between voxel and the the surface of PTV63
Distance to PTV70	Shortest path between voxel and the the surface of PTV70
Influence	Sum of influence matrix elements for the voxel

Table 4: The ten features used in the RF to predict the dose for any voxel.

The random forest used ten custom features outlined in Table 5 to predict the dose delivered to each voxel in the patient. The RF was trained with ten trees, and default settings with the `randomForestRegressor` from `scikit-learn`.