impact. The dosimetric impact of using different IVDTs when modifying energy, reconstruction kernel and patient size individually are below 2 %, for all Cyberknife and Tomotherapy plans considered. This is also the case for most of our pCT images. In extreme cases for tCT, e.g. when comparing a small patient scanned at 100 kV using the FC64 reconstruction kernel compared to a large patient scanned at 135 kV using the FC13 kernel, HU differences up to 900 (in bone) can be obtained leading to systematic dose differences up to 6 % (DVH shift). Using an "average" IVDT still leads to dose uncertainties > 2 %. Results can be CT scanner specific.

Conclusion

Uncertainties on pCT images used for MRI-only treatment planning should be compared to those on tCT images. The uncertainties on tCT images (even when not considering CT artifacts) are non-negligible and are of the same order as those on pCT images generated by e.g. atlas-based methods.

Electronic Poster: Physics track: (Quantitative) functional and biological imaging

EP-1677 Multicentre initiative for standardisation of image biomarkers

A. Zwanenburg[1], Image Biomarker Standardisation Initiative IBSI[2]

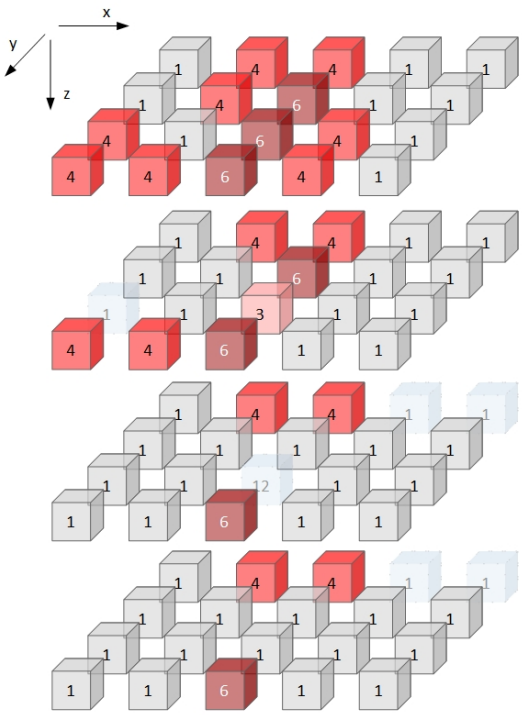
[1]*OncoRay - National Center for Radiation Research in Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus - Technische Universität Dresden - Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany*

Purpose or Objective

Personalised cancer treatment has the potential to improve patient treatment outcomes. One particular approach to personalised treatment is radiomics. Radiomics is the high-throughput analysis of medical images. There are several challenges within the radiomics field which need to be overcome to translate findings into clinical practice. The Image Biomarker Standardisation Initiative (IBSI) addresses the challenge of reproducing and validating reported findings by comparing and standardising definitions and implementation of several image feature sets between participating institutions.

Material and Methods

A 5x4x4 voxel digital phantom was devised, with a super-imposed region-of-interest (ROI) mask (Figure 1). This volume has characteristics similar to real patient volumes of interest, namely voxels outside of the ROI and missing grey levels. The phantom is moreover sufficiently small to manually calculate features for validation purposes. Because no pre-processing steps (e.g. discretisation) are necessary for calculations on the phantom, feature values may be standardised across all institutions.



A set of definitions for statistical, morphological and textural features was compiled. Commonly used texture matrices were included: the grey level co-occurrence matrix (GLCM), the run length matrix (GLRLM), the size zone matrix (GLSZM), the distance zone matrix (GLDZM), the neighbourhood grey tone difference matrix (NGTDM) and the neighbouring grey level dependence matrix (NGLDM). The definitions and the digital phantom were shared with all participating institutions. The participants then extracted image features from the phantom and reported them. Differences and similarities between participants were discussed to investigate potential errors and necessary changes made to achieve a standard value. Texture matrices can be evaluated per image slice (2D) or in a volume (3D). GLCM and GLRLM are moreover calculated for 4 (2D) or 13 (3D) directional vectors to achieve rotational invariance. GLCM and GLRLM features are then either calculated for every direction and averaged (avg), or after merging the matrices into a single matrix (mrg).

Results

17 features were standardised between institutions (Table 1). 58 features are close to standardisation, with one institution with a deviating value. The standardisation of the remaining features is ongoing.

Feature set		# features	# institutions	median CoV [min - max] (%)	# full agreement	# all but 1 agreement
Statistics		17	12	1 [0 - 256]	6	10
Morphology		15	12	48 [22 - 244]	0	0
GLCM	2D avg	25	4	62 [1 - 373]	0	0
	3D avg	25	8	42 [1 - 283]	0	2
	3D mrg	25	5	2 [0 – 31]	1	18
GLRLM	2D avg	16	3	44 [1 - 122]	0	-
	3D avg	16	8	34 [1 - 281]	0	2
	3D mrg	16	6	1 [0 - 158]	1	12
GLSZM	2D	16	3	54 [0 – 86]	3	-
	3D	16	10	13 [1 - 314]	0	6
GLDZM	2D	16	1	-	0	-
	3D	16	4	22 [0 - 200]	4	10
NGLDM	2D	16	1	-	0	0
	3D	16	5	16 [0 – 85]	2	15
NGTDM	2D	5	3	62 [56 - 127]	0	-
	3D	5	9	217 [122 - 300]	0	0

Conclusion
Definitions for a number of image features were devised and evaluated on a digital phantom within an international network. The feature definitions, digital phantom and corresponding feature values will be made available as a standard benchmark database for use by other institutions.

EP-1678 Are PET radiomic features robust enough with respect to tumor delineation uncertainties?
M.L. Belli[1], S. Broggi[1], C. Fiorino[1], V. Bettinardi[2], F. Fallanca[2], E.G. Vanoli[2], I. Dell'Oca[3], P. Passoni[3], N. Di Muzio[3], R. Calandrino[1], M. Picchio[2], G.M. Cattaneo[1]
[1]*San Raffaele Scientific Institute, Medical Physics, Milano, Italy*
[2]*San Raffaele Scientific Institute, Nuclear Medicine, Milano, Italy*
[3]*San Raffaele Scientific Institute, Radiotherapy, Milano, Italy*

Purpose or Objective
Radiomic techniques convert imaging data into a high dimensional feature space, guided by the hypothesis that these features may capture distinct tumor phenotypes predicting treatment outcome; it is clear that large multi Institutional studies are needed. The accuracy of tumor contouring based on PET is still a challenge issue in radiotherapy(RT) and this may strongly influence the extraction of radiomic parameters. Aim of current work was to investigate the robustness of PET radiomic features with respect to tumour delineation uncertainty in two clinically relevant situations.
Material and Methods

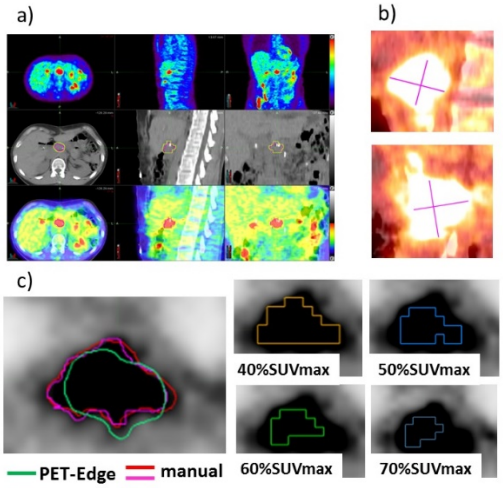


Figure 1

Twenty-five head-and-neck (HNC, with both T and N lesion) and twenty-five pancreatic (with only Tsite) cancer patients(pts) were considered. Patient images were acquired on three different PET/CT scanners with different characteristics and protocol acquisition. Seven contours were delineated for each lesion of the 50pts following different methods using the software MIM(Figure1.a): 2 different manual contours(Figure1.c) 1 semi-automatic ('PET-edge"based on maximum gradient detection, Figure1.b), and 4 automatic (based on a threshold:40%,50%,60%,70% of the SUVmax). The open access CGITAsoftware was used to extract several texture features (TA, e.g. entropy,skewness,dissimilarity,….) divided into different parent matrices (e.g. Co-occurrence,Voxel-alignment,…). Contours were compared in terms of both volume agreement (DICEindex) as well as TA difference (Kruskal-Wallis test). 9 manual contours were also blinded re-contoured, and the intra-observer variability was also evaluated (DICEindex). Furthermore, the repeatability of semi-automatic contouring was also tested.
Results
A total of 73 TA were extracted on each contour. A strong disagreement was found between automatic SUVmax threshold contours and manual or semi-automatic contours in terms of both DICE and TA agreement (9/73 TA for HNC and 10/73 for pancreas pts with p-value>0.05,Figure 2). Instead, both the inter-observer as well as the agreement between manual and semi-automatic contour was relatively high, for both volume (median DICE=0.71,range=0.36-0.96) and TA extraction (72/73 with p-value>0.05 for both HNC and pancreas pts). A high intra-observer agreement and a high contour repeatability were found for manual contours (median DICE=0.75,range:0.13-0.92) and for the semi-automatic method for lesions with high uptake values (median DICE=0.95,range=0.42-1.00). No statistically significant difference was found among scanners (p-value=0.12).

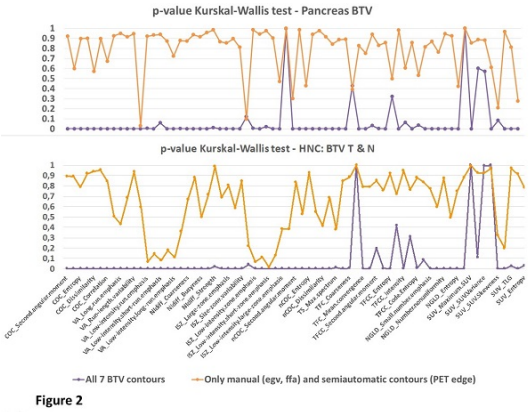


Figure 2

Conclusion
Almost the totality of the selected radiomic features were sufficiently robust against the delineation when using manual and semi-automatic methods, while threshold based methods resulted to be less robust. The satisfactory results with a semi-automatic PET contouring method suggests, for the two clinically situations considered in this work, possible promising applications for consistent and fast textural feature extraction in multi-centric studies.

EP-1679 Preliminary functional imaging study on an integrated 1.5T MR-Linac machine
M. Kadbi[1], Y. Ding[2], J. Wang[2], C.D. Fuller[3]
[1]*Philips, MR Therapy, Gainesville, USA*
[2]*MD Anderson, Department of Radiation Physics, Houston, USA*
[3]*MD Anderson, Department of Radiation Oncology, Houston, USA*

Purpose or Objective
Diffusion-weighted imaging (DWI) is a promising technique in MR guided radiotherapy (MRgRT) to delineate the tumor, predict response to induction chemotherapy, response to radiation therapy, and has been demonstrated as a biomarker of recurrence. This is the first attempt to investigate the performance of DWI technique in an integrated MR-Linac which combines Philips 1.5T MRI with 7 MV photon beam Elekta Linear accelerator (Linac). Conventional EPI-based DWI was compared with Spin-Echo (SE)-based DWI and geometrical distortion of the sequences were benchmarked with CT images as reference for geometric fidelity.