Impact of CT convolution kernel on robustness of radiomic features for different lung diseases and tissue types

Denzler, Sarah ; Vuong, Diem ; Bogowicz, Marta ; Pavic, Matea ; Frauenfelder, Thomas ; Thierstein, Sandra ; Eboulet, Eric Innocents ; Maurer, Britta ; Schniering, Janine ; Gabryś, Hubert Szymon ; Schmitt-Opitz, Isabelle ; Pless, Miklos ; Foerster, Robert ; Guckenberger, Matthias ; Tanadini-Lang, Stephanie

Abstract: **OBJECTIVES** In this study, we aimed to assess the impact of different CT reconstruction kernels on the stability of radiomic features and the transferability between different diseases and tissue types. Three lung diseases were evaluated, i.e. non-small cell lung cancer (NSCLC), malignant pleural mesothelioma (MPM) and interstitial lung disease related to systemic sclerosis (SSc-ILD) as well as four different tissue types, i.e. primary tumor, largest involved lymph node ipsilateral and contralateral lung. **METHODS** Pre-treatment non-contrast enhanced CT scans from 23 NSCLC, 10 MPM and 12 SSc-ILD patients were collected retrospectively. For each patient, CT scans were reconstructed using smooth and sharp kernel in filtered back projection. The regions of interest (ROIs) were contoured on the smooth kernel-based CT and transferred to the sharp kernel-based CT. The voxels were resized to the largest voxel dimension of each cohort. In total, 1386 features were analyzed. Feature stability was assessed using the intraclass correlation coefficient. Features above the stability threshold >0.9 were considered stable. **RESULTS** We observed a strong impact of the reconstruction method on stability of the features (at maximum 26% of the 1386 features were stable). Intensity features were the most stable followed by texture and wavelet features. The wavelet features showed a positive correlation between percentage of stable features and size of the ROI ($R^2 = 0.79$, $p = 0.005$). Lymph node radiomics showed poorest stability ($<10\%$) and lung radiomics the largest stability (26%). Robustness analysis done on the contralateral lung could to a large extent be transferred to the ipsilateral lung, and the overlap of stable lung features between different lung diseases was more than 50%. However, results of robustness studies cannot be transferred between tissue types, which was investigated in NSCLC and MPM patients; the overlap of stable features for lymph node and lung, as well as for primary tumor and lymph node was very small in both disease types. **CONCLUSION** The robustness of radiomic features is strongly affected by different reconstruction kernels. The effect is largely influenced by the tissue type and less by the disease type. **ADVANCES IN KNOWLEDGE** The study presents to our knowledge the most complete analysis on the impact of convolution kernel on the robustness of CT-based radiomics for four relevant tissue types in three different lung diseases. .

DOI: <https://doi.org/10.1259/bjr.20200947>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-200947>

Journal Article

Published Version

Originally published at:

Denzler, Sarah; Vuong, Diem; Bogowicz, Marta; Pavic, Matea; Frauenfelder, Thomas; Thierstein, Sandra; Eboulet, Eric Innocents; Maurer, Britta; Schniering, Janine; Gabryś, Hubert Szymon; Schmitt-Opitz, Isabelle; Pless, Miklos; Foerster, Robert; Guckenberger, Matthias; Tanadini-Lang, Stephanie (2021). Impact of CT convolution kernel on robustness of radiomic features for different lung diseases and tissue types. *British Journal of Radiology*, 94(1120):20200947.

DOI: <https://doi.org/10.1259/bjr.20200947>

Received:
05 August 2020

Revised:
03 December 2020

Accepted:
26 January 2021

<https://doi.org/10.1259/bjr.20200947>

Cite this article as:

Denzler S, Vuong D, Bogowicz M, Pavic M, Frauenfelder T, Thierstein S, et al. Impact of CT convolution kernel on robustness of radiomic features for different lung diseases and tissue types. *Br J Radiol* 2021; **94**: 20200947.

FULL PAPER

Impact of CT convolution kernel on robustness of radiomic features for different lung diseases and tissue types

¹SARAH DENZLER, ¹DIEM VUONG, ¹MARTA BOGOWICZ, PhD, ¹MATEA PAVIC, MD, ²THOMAS FRAUENFELDER, MD, SANDRA THIERSTEIN, ³ERIC INNOCENTS EBOULET, ⁴BRITTA MAURER, PD MD, ⁴JANINE SCHNIERING, PhD, ¹HUBERT SZYMON GABRYŚ, PhD, ⁵ISABELLE SCHMITT-OPITZ, MD, ⁶MIKLOS PLESS, MD, ¹ROBERT FOERSTER, PD MD, ¹MATTHIAS GUCKENBERGER, MD and ¹STEPHANIE TANADINI-LANG, PhD

¹Department of Radiation Oncology, University Hospital Zurich and University of Zurich, Zurich, Switzerland

²Institute of Diagnostic and Interventional Radiology, University Hospital Zurich and University of Zurich, Zurich, Switzerland

³Swiss Group for Clinical Cancer Research (SAKK) Coordinating Center, Bern, Switzerland

⁴Department of Rheumatology, Center of Experimental Rheumatology, University Hospital Zurich and University of Zurich, Zurich, Switzerland

⁵Department of Thoracic Surgery, University Hospital Zurich and University of Zurich, Zurich, Switzerland

⁶Department of Medical Oncology, Kantonsspital Winterthur, Winterthur, Switzerland

Address correspondence to: Sarah Denzler

E-mail: sarah_denzler@hotmail.com; sarah.denzler@gzo.ch

Objectives: In this study, we aimed to assess the impact of different CT reconstruction kernels on the stability of radiomic features and the transferability between different diseases and tissue types. Three lung diseases were evaluated, *i.e.* non-small cell lung cancer (NSCLC), malignant pleural mesothelioma (MPM) and interstitial lung disease related to systemic sclerosis (SSc-ILD) as well as four different tissue types, *i.e.* primary tumor, largest involved lymph node ipsilateral and contralateral lung.

Methods: Pre-treatment non-contrast enhanced CT scans from 23 NSCLC, 10 MPM and 12 SSc-ILD patients were collected retrospectively. For each patient, CT scans were reconstructed using smooth and sharp kernel in filtered back projection. The regions of interest (ROIs) were contoured on the smooth kernel-based CT and transferred to the sharp kernel-based CT. The voxels were resized to the largest voxel dimension of each cohort. In total, 1386 features were analyzed. Feature stability was assessed using the intraclass correlation coefficient. Features above the stability threshold >0.9 were considered stable.

Results: We observed a strong impact of the reconstruction method on stability of the features (at maximum 26% of the 1386 features were stable). Intensity features

were the most stable followed by texture and wavelet features. The wavelet features showed a positive correlation between percentage of stable features and size of the ROI ($R^2 = 0.79$, $p = 0.005$). Lymph node radiomics showed poorest stability (<10%) and lung radiomics the largest stability (26%). Robustness analysis done on the contralateral lung could to a large extent be transferred to the ipsilateral lung, and the overlap of stable lung features between different lung diseases was more than 50%. However, results of robustness studies cannot be transferred between tissue types, which was investigated in NSCLC and MPM patients; the overlap of stable features for lymph node and lung, as well as for primary tumor and lymph node was very small in both disease types.

Conclusion: The robustness of radiomic features is strongly affected by different reconstruction kernels. The effect is largely influenced by the tissue type and less by the disease type.

Advances in knowledge: The study presents to our knowledge the most complete analysis on the impact of convolution kernel on the robustness of CT-based radiomics for four relevant tissue types in three different lung diseases.

INTRODUCTION

Malfunction of the lung has a strong impact on the quality of life. Patients with lung diseases frequently suffer from dyspnea, coughing and general weakness. Various diseases

originate from the lung tissue. Lung cancer is one of the leading causes of cancer worldwide and the leading cause for cancer-related death,¹ with non-small cell lung cancer (NSCLC) being the most common type.^{2,3} Malignant

pleural mesothelioma (MPM) is a rare but even more aggressive cancer with a dismal prognosis.⁴ Besides cancer, there are also other forms of lung diseases with strong impact on the patient's life and well-being. For example, interstitial lung disease is the leading cause of death in patients with systemic sclerosis (SSc-ILD) and is characterized by various degrees of lung inflammation and fibrosis.

Various options to treat these lung diseases exist, however, the selection of the right treatment for the right patient is challenging. There is consequently a strong clinical need for decision-support systems to guide the optimal treatment.⁵ Biomarker identification has increasingly gained importance in precision medicine to deliver the optimal treatment to every individual patient. Quantitative imaging biomarkers were proposed to provide an observer-independent, non-invasive and comprehensive characterization of the region of interest (ROI). Medical imaging technologies are increasingly used to identify prognostic and predictive biomarkers for decision-support systems⁶; they can provide a 3D profile without the risk of sampling errors, allow assessment with high temporal resolution, are acquired non-invasively and use existing resources for data acquisition, analysis and storage. The extraction of quantitative image features is also referred to as radiomics⁷ and can be classified in shape, intensity, texture and filtered-based features.⁸ Radiomics has been successfully applied to characterize lung cancer and lung fibrosis^{9–11} and several CT radiomics studies have shown that these markers could be linked to treatment outcome.^{12,13}

Despite their potential prognostic power, radiomic features show high scanner and image acquisition dependency,¹⁴ including image reconstruction, intra- and interobserver variability in image segmentation, respiratory and cardiac motion during imaging process and other random effects,^{15–17} hampering the clinical implementation. Several studies have investigated the feature stability against reconstruction techniques in different cancer types^{18–20} and have shown the choice of convolution kernel in filtered back projection (FBP) CT reconstruction to be of major impact. Convolution kernel is often applied on the raw projection data to account for the limitation of the conventional back-propagation algorithm which tends to provide blurred images. Different convolution kernels are available to additionally exploit and enhance the underlying information of the specific tissue. For lung imaging examination, smooth and sharp convolution kernel are the most often used. A smooth kernel aims to suppress high frequency domain information to reduce noise and leads to more smoothed images. A sharp kernel enhances structure heterogeneity leading to more clearly visible brightness details or textures of the lung structures, such as small vessels and bronchi. However, these images do contain more noise. Different groups have investigated the robustness of radiomic features for the two most often used convolution kernels for FBP in chest CT imaging.^{21,22} They concluded that reconstruction method and slice thickness significantly affected different features. Their phantom-based study evaluated the robustness of a low number of features (14 features) and focused on the primary tumor only.

In this study, we aimed to assess the robustness of radiomic features in two different CT reconstruction kernels (smooth and sharp kernel) for three different lung diseases: NSCLC, MPM as well as for SSc-ILD patients. Further, a complete robustness analysis of the relevant tissue types (primary tumor, lymph node, ipsilateral and contralateral lung) was performed. Finally, we assessed if robustness studies performed for a specific tissue and disease type can be transferred to other disease or tissue types.

METHODS AND MATERIALS

Studied cohorts of patients

45 patients of 3 different lung disease were included. The first cohort (denoted as NSCLC) consisted of 23 retrospectively collected patients with locally advanced Stage IIIA/N2 NSCLC cancer (sixth TNM edition) from a Swiss multicentric prospective Phase III randomized trial (SAKK 16/00²³). The second cohort consisted of 10 Stage III MPM patients treated at the University Hospital Zurich (USZ).²⁴ The third cohort consisted of 12 patients with proven systemic sclerosis and mild stage of fibrosis (denoted as SSc-ILD cohort) treated at the USZ. MPM and SSc-ILD patients were retrospectively collected from our institution University Hospital Zurich, (USZ). Written consent was obtained from all patients and this study was accepted by the local ethical committee (KEK ZH PB_2018–0026).

Image acquisition

For each patient with oncological disease, a native pre-treatment CT was collected as a part of standard PET/CT examination. For the SSc-ILD cohort, a high-resolution CT was collected. The images were reconstructed using FBP once with a sharp and once with a smooth kernel from the same raw data. As it was a retrospective study, we assessed the kernels that were available, which were standard and lung kernel. The phantom study of Mackin et al²⁵ compared which kernel of a different vendor matched closest to GE standard kernel and showed that Siemens B30f had the closest match, *i.e.* features were the most similar to the features calculation on a CT using GE standard kernel. They further stated that B60f was most similar to the GE lung kernel. Our selection for I70f was also based on the availability of the kernel. Detailed information on the CT images can be found in Table 1.

Delineation

We considered for each patient at maximum four tissue types: primary tumor (GTV_prim), largest lymph node involved (GTV_LN_max), ipsilateral (lung_ipsi) and contralateral lung (lung_contra). In SSc-ILD patients, only lung_ipsi and lung_contra were considered. All tissue types were delineated on the smooth kernel-based CT and contours were transferred to the corresponding registered sharp kernel based CT. GTV_prim and GTV_LN_max were delineated manually, lung_ipsi and lung_contra using an automatic threshold segmentation. The NSCLC cohort was contoured according to the institutional protocol for routine radiotherapy practice.²⁶ For the MPM cohort, a special study protocol was defined where primary tumor comprised pleural thickening as well as lung nodules, infiltrated pericard and mediastinal extensions. For NSCLC and MPM, the lung and mediastinal windows were set for the delineation and the corresponding [¹⁸F]-FDG PET scans were used to assist in

Table 1. Overview of image acquisition characteristics

Type	NSCLC		MPM		SSc-ILD	
Manufacturer	GE Medical Systems		GE Medical Systems		SIEMENS	
Model	Discovery STE, Discovery 690		Discovery 690		SOMATOM Definition AS	
Reconstruction	Smooth	Sharp	Smooth	Sharp	Smooth	Sharp
	STANDARD	LUNG	STANDARD	DETAIL	B30f	I70f
Resolution (mm)	0.98–1.37 × 0.98–1.37 × 1.25–3.27	0.94–1.37 × 0.94–1.37 × 1.25–3.27	0.98 × 0.98 × 3.27, 1.17 × 1.17 × 3.27	0.59–0.76 × 0.59–0.76 × 1.25	0.62–0.85 × 0.62–0.85 × 0.8–1.6	0.54–0.84 × 0.54–0.84 × 0.8–1.0

MPM, malignant pleural mesothelioma; NSCLC, non-small cell lung cancer; SSc-ILD, interstitial lung disease related to systemic sclerosis.

NSCLC cohort was retrospectively collected from a multi centric clinical trial and a large variation with in image spatial resolution can be observed in contrast to MPM and SSc-ILD which were imaged in one single institution.

contouring. Pleural effusion and atelectasis were excluded from lung, tumor and lymph node contours. For the contouring of the lungs, a semi-automated method based on the HU thresholding was used. In SSc-ILD, the right lung was used for further analysis. Delineation was performed in MIMVista (MIM software Inc., v. 6.7.9, Cleveland, OH). An example of the delineation can be found in Figure 1.

Radiomics analysis and image pre-processing

For the radiomics calculations, we used an in-house developed software (Z-Rad) written in Python programming language (v. 2.7.14.). In total, 1386 features were extracted:

- Intensity distribution ($n = 17$).

Figure 1. CT axial slices for three patients showing the delineated tissue types in the two kernel CT. NSCLC in smooth (a) and in sharp kernel (b), MPM in smooth (c) and in sharp kernel (d), SSc-ILD in smooth (e) and in sharp kernel (f). Right lung (green), left lung (blue), GTV_prim (red) and LN_max (pink) are shown. MPM, malignant pleural mesothelioma; NSCLC, non-small cell lung cancer; SSc-ILD, interstitial lung disease related to systemic sclerosis.

- Texture ($n = 37$): the Gray Level Co-occurrence Matrix ($n = 52$), the Neighborhood Gray Tone Difference Matrix ($n = 5$), the Gray Level Run Length Matrix ($n = 32$), the Gray Level Size Zone Matrix ($n = 16$) and the Neighboring Gray Level Dependence Matrix ($n = 16$).
- Wavelet transform ($n = 1232$)

No shape features were analyzed since a high stability was assumed as the same contours were used in both image sets. Definition of intensity and texture features were standardized according to the image biomarker standardization initiative (IBSI, v. 11).²⁷ For further details on the feature calculations, consult <https://medical-physics-usz.github.io>.

Due to the retrospective nature of this study, image resolution differed between the patients. Therefore, CT images of both convolution kernels were resized to cubic voxels of 3.27 mm for NSCLC and 3.27 mm for MPM and 1.6 mm for SSc-ILD using linear interpolation. These voxel sizes correspond to the lowest image resolution in the CT scans in the respective cohort. A bin size of 10 HU for all texture and wavelet analyses was chosen. For the analysis of the primary tumor and the lymph nodes, a Hounsfield unit (HU) range of -300 to 200 was set to exclude bone, air and lung tissue. For the lung analysis, a HU range of -1024 to 50 was set to exclude non-lung tissue structures.

Statistical analysis was performed using R (v. 3.3.2.). Intraclass correlation coefficient (ICC) was calculated according to Shrout et al.²⁸

$$ICC(3,1) = \frac{BMS - EMS}{BMS + (\kappa - 1)EMS}$$

where BMS is the between subject mean squares and EMS is the residual mean squares from Friedman's two-way ANOVA. The variable κ denotes the number of observers, *i.e.* the number of repeated measurements (here, scans with different convolution kernels). For both investigations, an ICC larger than 0.9 was considered stable among both types of CT scans. Despite the small patient cohort, the high acceptance level of the ICC kept Type I and Type II errors small.²⁹

RESULTS

Delineation

Table 2 shows median volume and HU ranges for all three lung diseases. The relative range of volume was largest for GTV_prim for both NSCLC and MPM. Accordingly, there was a large range for lung_ipsi because tumorous tissue was not included in the lung_ipsi. Primary tumors had more homogeneous HU compared to LN structures.

Stability of radiomic features within same disease

Overall, there was a large influence of the convolution kernel on the robustness of radiomic features (Table 3). The percentage of stable features was low ($<26\%$) and strongly depended on the tissue type. 3% of features were stable for GTV_LN_max in MPM, and the highest stability (26%) was observed for lung_ipsi and GTV_prim in MPM. In general, high stability was observed for lung tissue and GTV_prim, and low stability for

Table 2. Median volume (range) of structures for all tumor entities in ml and the absolute median relative HU difference between the different convolution kernels for different ROI are listed

	NSCLC			Mesothelioma			SSc-ILD		
	Volume median (min-max) (ml)	HU _{min} (%)	HU _{max} (%)	Volume median (min-max) (ml)	HU _{min} (%)	HU _{max} (%)	Volume median (min-max) (ml)	HU _{min} (%)	HU _{max} (%)
GTV_prim	25.51 (0.33–190.60)	0	3.11	566.033 (3.10–1866.75)	38.28	0		–	
GTV_LN_max	7.05 (1.41–30.93)	22.54	18.02	2.21 (0.94–12.37)	67.95	72.43	–	–	
Lung_ipsi	1915.38 (1290.13–3445.16)	3.15	0.59	968.27 (638.26–2488.83)	0	0	2701.01 (1597.37–3356.34)	0	0
Lung_contra	2023.88 (1263.76–3154.42)	0	0.05	1923.11 (877.96–2855.19)	0	0			

HU, Hounsfield unit; NSCLC, non-small cell lung cancer; ROI, region of interest; SSc-ILD, interstitial lung disease related to systemic sclerosis.

Largest HU difference was observed for the lymph node (GTV_LN_max). A smaller HU difference was observed for larger ROI, in particular for lung ROI where a HU limit was set.

Table 3. Percentages of stable features in NSCLC, MPM and SSc-ILD for different tissue types

Lung disease	Tissue type	Stable features (%)
NSCLC	GTV_prim	23%
	GTV_LN_max	8%
	Lung_ipsi	20%
	Lung_contra	19%
MPM	GTV_prim	26%
	GTV_LN_max	3%
	Lung_ipsi	26%
	Lung_contra	19%
SSc-ILD	Lung	19%

MPM, malignant pleural mesothelioma; NSCLC, non-small cell lung cancer; SSc-ILD, interstitial lung disease related to systemic sclerosis. Overall, the stability of the features is poor for all disease types. Lung tissue showed highest stability and lymph nodes radiomics the poorest stability.

the GTV_LN_max. A positive correlation was observed for the subset of intensity features, but not other feature types ($R^2=0.79$, $p=0.005$, [Supplementary Material 1](#)).

The ipsi- and contralateral lung had the highest percentage of stable features in intensity, texture and wavelet compared to the other studied tissue types. Lowest feature stability was observed for GTV_LN_max. The GTV_prim structure had more robust features than GTV_LN_max in MPM and NSCLC. Wavelet features had the highest percentage of stable features for GTV_prim (27.11%). Stable features of lung_ipsi and lung_contra had high agreement among all feature types. In SSc-ILD, there was the highest percentage of stable features in intensity features, followed by texture and wavelet features ([Figure 2](#)). A complete list of all ICC values of each feature can be found in [Supplementary Material 2](#).

Transferability of stable features across tissue types

The low feature stability rate in tissue types translate into a low number of features transferrable across the tissue types for both NSCLC (39 features, 2.8%) and MPM (5 features, 0.4%). Considering excluding the tissue type with the lowest stability rate, *i.e.* GTV_LN_max the overlap increased to 130 (9.4%) and 105 (7.6%) features for NSCLC and MPM, respectively ([Figure 3](#)).

Interestingly to be observed was the overlap of ipsi- and contralateral lung radiomics in NSCLC and MPM. There were 208 (15%) features that were stable in the ipsi- and the contralateral lung in NSCLC and MPM. Overall, in the ipsilateral lung there were more stable features than in the contralateral lung ([Figure 4](#)). In NSCLC, 75% of the features that were stable in the ipsilateral lung were also stable in the contralateral lung, for MPM this overlap was smaller (57.6%).

Figure 2. Percentage of stable features for NSCLC, MPM and SSc-ILD patients subdivided into feature types. Results of SSc-ILD are shown for one lung. MPM, malignant pleural mesothelioma; NSCLC, non-small cell lung cancer; SSc-ILD, interstitial lung disease related to systemic sclerosis.

Transferability of stable features across disease types

The observed results on the same disease type could only be transferred to a limited extend across the lung diseases ([Figure 5](#)). The overlap of stable features in GTV_prim, lung_ipsi and lung_contra was low, *i.e.* in GTV_prim there were 195 features (14%), in lung_ipsi 217 (15.7%) and in lung contra 130 (9.4%) shared by both NSCLC and MPM. No transferability of stable features was possible for GTV_LN_max tissue.

In particular, interesting was the transferability of stable features of lung radiomic features of all three lung diseases studied. There were 156 (11.3%) features that were stable over all 3 entities, which was around half of the features defined as stable based on one disease ([Figure 6](#)). NSCLC cohort had 277 stable features, 172 (12.4%) of them were also stable in SSc-ILD and 217 (15.7%) in Mesothelioma. In SSc-ILD, there were 269 stable features from which 179 (12.9%) were also stable for Mesothelioma.

DISCUSSION

The robustness of CT radiomic features in respect to reconstruction with different kernels was assessed in different lung diseases and tissue types. Radiomic features were strongly affected by differences in convolution kernel. For intensity features, there was a correlation between feature stability and size of the analyzed ROI – larger ROIs had a higher percentage of stable features. This

Figure 3. Overlap of stable radiomic features across tissue types in NSCLC and MPM. There was an overlap of features that were stable in GTV_prim, GTV_LN_max and lung_contra, but only a small number of features. More features were stable in all three tissue types in NSCLC than in MPM. More features were stable in GTV_prim and lung_contra than GTV_LN_max and other tissue types. MPM, malignant pleural mesothelioma; NSCLC, non-small cell lung cancer; SSC-ILD, interstitial lung disease related to systemic sclerosis.

Overlap of stable radiomic features across tissue types

might be due to the differences in slice thickness between NSCLC and MPM patients. There was no such correlation for the other feature types (wavelet and texture). Robustness analysis done on the contralateral lung could to a large extent be transferred to the ipsilateral lung and the overlap of stable lung features between different lung diseases was more than 50%. However, results of robustness studies cannot be transferred between tissue types, which was investigated in NSCLC and MPM patients; the overlap of stable features for LN and lung, as well as for primary tumor and LN was very small in both disease types. Comparing the stable

Figure 4. Overlap of stable features in the ipsi- and the contralateral lung in NSCLC and MPM. There were more stable features in the ipsilateral lung than in the contralateral lung. A large number of the stable features were transferrable for both lung sides. MPM, malignant pleural mesothelioma; NSCLC, non-small cell lung cancer.

Overlap of stable lung radiomic features

Figure 5. Overlap of stable radiomic features for NSCLC and MPM by tissue types. There was an overlap of stable features in all tissue types but GTV_LN_max. Lung_ipsi had the largest number of stable feature overlap in NSCLC and MPM. MPM, malignant pleural mesothelioma; NSCLC, non-small cell lung cancer.

Overlap of stable radiomic features stratified by tissue types

features across the three different tissue types in the same disease, only a small fraction of features overlapped (39 and 5 features in NSCLC and MPM).

These results challenge phantom studies to assess robustness of features with changing reconstruction kernel.^{30,31} Phantom studies have the advantage that they can evaluate the robustness of many parameters at the same time as well as they can assess a larger range of parameters such as large changes in voltage or current. However, in the light of our study one should be critical in directly transferring these results to the patient.

Several studies have shown that the reconstruction algorithm has a large influence on the stability of radiomic features. Two studies investigated the influence for lung cancer in specific. Zhao et al evaluated 89 features (intensity, texture and wavelets) and found that only approximately 10% of the features were stable in respect to changing the reconstruction kernel.³² In contrast, Kim et al found 40% of the features to be stable in respect to reconstruction algorithm.³³ However, only 15 features were assessed. We assessed in total 1386 features (intensity, texture and wavelet) and found 23% of them to be stable for NSCLC and 26% for MPM. The percentage of stable features is consistent between cancer types and therefore it is likely that the stability differences between our study and previous studies are due to a reduced number of features in previous studies.

Figure 6. Overlap of stable features for NSCLC, Mesothelioma and SSc-ILD. 11.3% of the features were stable in NSCLC, MPM and SSc-ILD. Both lung_contra and lung_ipsi radiomics are considered. MPM, malignant pleural mesothelioma; NSCLC, non-small cell lung cancer; SSc-ILD, interstitial lung disease related to systemic sclerosis.

Overlap of stable lung radiomic features across disease types

The overlap of the stable features between ipsilateral and contralateral lung was very high for both, NSCLC and MPM. However, this overlap was reduced when comparing the contralateral lung of NSCLC and MPM. Next to different imaging protocols and scanners differences in underlying pulmonary disease, *e.g.* smoking history and COPD could potentially explain our observation. The overlap is further reduced when also considering the lung tissue of SSc-ILD patients. This might be due to the different motion management (breath-hold for lung fibrosis, free breathing for NSCLC and MPM) and the significantly changed lung structure in the case of lung fibrosis.

A limitation of our study is the different image protocols used in different hospitals for NSCLC, which could also be a strength as it reflects real-life conditions with a variety of CT scanners used in different hospitals. In particular, definition of sharp and smooth kernel can substantially differ between CT scanners. Even though the imaging protocols were identical for MPM and SSc-ILD patients, the results of stable features transferability across these diseases may have been influenced by scanner variability. An additional limitation that applies to all robustness studies is that robust features might not correlate well with treatment outcome. He et al³⁴ have shown that a model based on standardized data performs better compared to a model based on non-standardized data for lung cancer. This suggests that standardization is extremely important for the future and can only partially be compensated by building robust models. If correction algorithms improves stability of features as suggested by Shafiq-Ul-Hassan et al has to further explored in a follow-up study.³⁵

Another limitation is the choice of the resampling method. Up- and downsampling can be performed to resample the medical images prior to radiomics calculation. To our current knowledge, there is no consensus which is the optimal one. Downsampling, *i.e.* resample images to a lower resolution as performed in this study, will cause smoothing of images and hence potentially relevant information is lost. This is in particular important when comparing two image sets with different sharpness level, therefore this choice may likely impact and lead to more optimistic robustness results in contrast to upsampling approach. However, to resample images to a higher resolution will introduce undesired artificial information.^{36,37} Further, we have decided to use linear interpolation as it was the most common used method at the time of the study.

A way to alleviate differences between CT-kernels could be post-hoc batch normalization as described from Orlhac et al³⁸. They showed that ComBat removed protocol effects (including standard and lung kernel) on texture features. An advantage is that it is fast and easily available. However, this method can only be used given that there are more than 20 patient per protocol, which with respect to a multicentric trial is not always the case.

In clinical situations, the CT image data set is very heterogenous, to be able to use these images for radiomics models for overall survival prediction, we suggest using a smaller subset of the images of the cohort with homogenous image settings (CT data using the same reconstruction kernel) or only use stable radiomics features. Vuong et al compared robust to standardized CT radiomics models to predict overall survival for NSCLC patients. They showed that the model with robust features performed equally to the standardized model, but indicated that harmonizing the image data set should be favored.²⁶

In this study, we only focused on the impact of a reconstruction kernel, because previous studies have shown that it has the largest impact beside a voxel size that can be standardized.³⁰ This allowed us to do a comprehensive comparison between different tissue types and diseases. However, to transfer these results to future outcome studies other factors, such as delineation uncertainty,²⁴ motion³⁹ and scanner parameters,³⁶ have also to be considered.

Our findings showed that radiomic feature stability for non-contrast enhanced CT imaging in the lung is tissue- and disease-specific. A transfer from one disease to another is only possible for large structures that have similar characteristics and no transfer is possible between tissue types. This needs to be kept in mind when interpreting phantom studies on robustness of radiomic features.

CONCLUSION

Our study is, to our knowledge, the first to assess stability of features depending on reconstruction method in different tissue and lung disease types looking at over 1300 features. CT convolution kernel had a large impact on radiomic features. The robustness of radiomic features is strongly dependent on the tissue type and less-dependent on the disease type. This needs to be considered in further radiomic multicentric studies to develop robust radiomic models.

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394–424. doi: <https://doi.org/10.3322/caac.21492>
- Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clin Chest Med* 2011; **32**: 605–44. doi: <https://doi.org/10.1016/j.ccm.2011.09.001>
- van Baardwijk A, Wanders S, Boersma L, Borger J, Ollers M, Dingemans A-MC, et al. Mature results of an individualized radiation dose prescription study based on normal tissue constraints in stages I to III non-small-cell lung cancer. *J Clin Oncol* 2010; **28**: 1380–6. doi: <https://doi.org/10.1200/JCO.2009.24.7221>
- Yang H, Testa JR, Carbone M. Mesothelioma epidemiology, carcinogenesis, and pathogenesis. *Curr Treat Options Oncol* 2008; **9**(2-3): 147–57. doi: <https://doi.org/10.1007/s11864-008-0067-z>
- Ley B, Collard HR. Epidemiology of idiopathic pulmonary fibrosis. *Clin Epidemiol* 2013; **5**: 483–92. doi: <https://doi.org/10.2147/CLEP.S54815>
- Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014; **5**: 4006. doi: <https://doi.org/10.1038/ncomms5006>
- Bogowicz M, Vuong D, Huellner MW, Pavic M, Andratschke N, Gabrys HS, et al. Ct radiomics and PET radiomics: ready for clinical implementation? *Q J Nucl Med Mol Imaging* 2019; **63**: 355–370. doi: <https://doi.org/10.23736/S1824-4785.19.03192-3>
- Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012; **48**: 441–6. doi: <https://doi.org/10.1016/j.ejca.2011.11.036>
- Thawani R, McLane M, Beig N, Ghose S, Prasanna P, Velcheti V, et al. Radiomics and radiogenomics in lung cancer: a review for the clinician. *Lung Cancer* 2018; **115**: 34–41. doi: <https://doi.org/10.1016/j.lungcan.2017.10.015>
- Armato SG, Francis RJ, Katz SI, Ak G, Opitz I, Gudmundsson E, et al. Imaging in pleural mesothelioma: a review of the 14th International Conference of the International mesothelioma interest group. *Lung Cancer* 2019; **130**: 108–14. doi: <https://doi.org/10.1016/j.lungcan.2018.11.033>
- Obert M, Kampschulte M, Limburg R, Barańczuk S, Krombach GA. Quantitative computed tomography applied to interstitial lung diseases. *Eur J Radiol* 2018; **100**: 99–107. doi: <https://doi.org/10.1016/j.ejrad.2018.01.018>
- O'Connor JPB, Aboagye EO, Adams JE, Aerts HJWL, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 2017; **14**: 169–86. doi: <https://doi.org/10.1038/nrclinonc.2016.162>
- Lee YH, Bang H, Kim DJ. How to establish clinical prediction models. *Endocrinol Metab* 2016; **31**: 38–44. doi: <https://doi.org/10.3803/EnM.2016.31.1.38>
- Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012; **30**: 1234–48. doi: <https://doi.org/10.1016/j.mri.2012.06.010>
- Yip SSE, Aerts HJWL. Applications and limitations of radiomics. *Phys Med Biol* 2016; **61**: R150–66. doi: <https://doi.org/10.1088/0031-9155/61/13/R150>
- Larue RTHM, Defraene G, De Ruyscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol* 2017; **90**: 20160665. doi: <https://doi.org/10.1259/bjr.20160665>
- Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present... any future? *Eur J Nucl Med Mol Imaging* 2017; **44**: 151–65. doi: <https://doi.org/10.1007/s00259-016-3427-0>
- van Velden FHP, Kramer GM, Frings V, Nissen IA, Mulder ER, de Langen AJ, et al. Repeatability of radiomic features in non-small-cell lung cancer [(18)F]FDG-PET/CT studies: impact of reconstruction and delineation. *Mol Imaging Biol* 2016; **18**: 788–95. doi: <https://doi.org/10.1007/s11307-016-0940-2>
- Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol* 2010; **49**: 1012–6. doi: <https://doi.org/10.3109/0284186X.2010.498437>
- Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of intratumour ¹⁸F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging* 2013; **40**: 1662–71. doi: <https://doi.org/10.1007/s00259-013-2486-8>
- Zhao B, Tan Y, Tsai WY, Schwartz LH, Lu L. Exploring variability in CT characterization of tumors: a preliminary phantom study. *Transl Oncol* 2014; **7**: 88–93. doi: <https://doi.org/10.1593/tlo.13865>
- Milanese G, Mannil M, Martini K, Maurer B, Alkadhi H, Frauenfelder T. Quantitative CT texture analysis for diagnosing systemic sclerosis: effect of iterative reconstructions and radiation doses. *Medicine* 2019; **98**: e16423. doi: <https://doi.org/10.1097/MD.00000000000016423>
- Pless M, Stupp R, Ris H-B, Stahel RA, Weder W, Thierstein S, et al. Induction chemoradiation in stage IIIA/N2 non-small-cell lung cancer: a phase 3 randomised trial. *Lancet* 2015; **386**: 1049–56. doi: [https://doi.org/10.1016/S0140-6736\(15\)60294-X](https://doi.org/10.1016/S0140-6736(15)60294-X)
- Pavic M, Bogowicz M, Würms X, Glatz S, Finazzi T, Riesterer O, et al. Influence of inter-observer delineation variability on radiomics stability in different tumor sites. *Acta Oncol* 2018; **57**: 1070–4. doi: <https://doi.org/10.1080/0284186X.2018.1445283>
- Mackin D, Ger R, Gay S, Dodge C, Zhang L, Yang J, et al. Matching and Homogenizing convolution kernels for quantitative studies in computed tomography. *Invest Radiol* 2019; **54**: 288–95. doi: <https://doi.org/10.1097/RLI.0000000000000540>
- Vuong D, Bogowicz M, Denzler S, Oliveira C, Foerster R, Amstutz F, et al. Comparison of robust to standardized CT radiomics models to predict overall survival for non-small cell lung cancer patients. *Med Phys* 2020; **47**: 4045–53. doi: <https://doi.org/10.1002/mp.14224>
- Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative Radiomics for high-throughput image-based phenotyping. *Radiology* 2020; **295**: 328–38. doi: <https://doi.org/10.1148/radiol.2020191145>
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420–8. doi: <https://doi.org/10.1037/0033-2909.86.2.420>
- Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998; **17**: 101–10. doi: [https://doi.org/10.1002/\(SICI\)1097-0258\(19980115\)17:1<101::AID-SIM727>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E)
- Berenguer R, Pastor-Juan MDR, Canales-Vázquez J, Castro-García M, Villas

- MV, Mansilla Legorburo F, et al. Radiomics of CT features may be Nonreproducible and redundant: influence of CT acquisition parameters. *Radiology* 2018; **288**: 407–15. doi: <https://doi.org/10.1148/radiol.2018172361>
31. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* 2017; **44**: 1050–62. doi: <https://doi.org/10.1002/mp.12123>
 32. Zhao B, Tan Y, Tsai W-Y, Qi J, Xie C, Lu L, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep* 2016; **6**: 23428. doi: <https://doi.org/10.1038/srep23428>
 33. Kim H, Park CM, Lee M, Park SJ, Song YS, Lee JH, et al. Impact of reconstruction algorithms on CT radiomic features of pulmonary tumors: analysis of intra- and inter-reader variability and inter-reconstruction algorithm variability. *PLoS One* 2016; **11**: e0164924. doi: <https://doi.org/10.1371/journal.pone.0164924>
 34. He L, Huang Y, Ma Z, Liang C, Liang C, Liu Z. Effects of contrast-enhancement, reconstruction slice thickness and convolution kernel on the diagnostic performance of radiomics signature in solitary pulmonary nodule. *Sci Rep* 2016; **6**: 34921. doi: <https://doi.org/10.1038/srep34921>
 35. Shafiq-Ul-Hassan M, Zhang GG, Hunt DC, Latifi K, Ullah G, Gillies RJ, et al. Accounting for reconstruction kernel-induced variability in CT radiomic features using noise power spectra. *J Med Imaging* 2018; **5**: 011013. doi: <https://doi.org/10.1117/1.JMI.5.1.011013>
 36. Mackin D, Fave X, Zhang L, Yang J, Jones AK, Ng CS, et al. Harmonizing the pixel size in retrospective computed tomography radiomics studies. *PLoS One* 2017; **12**: e0178524. doi: <https://doi.org/10.1371/journal.pone.0178524>
 37. Zwanenburg A, Leger S, Agolli L, Pilz K, Troost EGC, Richter C, et al. Assessing robustness of radiomic features by image perturbation. *Sci Rep* 2019; **9**: 614. doi: <https://doi.org/10.1038/s41598-018-36938-4>
 38. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I. Validation of a method to compensate multicenter effects affecting CT Radiomics. *Radiology* 2019; **291**: 53–9. doi: <https://doi.org/10.1148/radiol.2019182023>
 39. Vuong D, Tanadini-Lang S, Huellner MW, Veit-Haibach P, Unkelbach J, Andratschke N, et al. Interchangeability of radiomic features between [18F]-FDG PET/CT and [18F]-FDG PET/MR. *Med Phys* 2019; **46**: 1677–85. doi: <https://doi.org/10.1002/mp.13422>