

Article

# CT Reconstruction Kernels and the Effect of Pre- and Post-Processing on the Reproducibility of Handcrafted Radiomic Features

Turkey Refae<sup>1,2,\*</sup>, Zohaib Salahuddin<sup>1,†</sup>, Yousif Widaatalla<sup>1,†</sup>, Sergey Primakov<sup>1,3</sup>, Henry C. Woodruff<sup>1,3</sup>, Roland Hustinx<sup>4</sup>, Felix M. Mottaghy<sup>3,5</sup>, Abdalla Ibrahim<sup>1,6,‡</sup> and Philippe Lambin<sup>1,3,‡</sup>

<sup>1</sup> The D-Lab, Department of Precision Medicine, GROW-School for Oncology, Maastricht University, 6200 Maastricht, The Netherlands; z.salahuddin@maastrichtuniversity.nl (Z.S.); y.widaatalla@maastrichtuniversity.nl (Y.W.); s.primakov@maastrichtuniversity.nl (S.P.); h.woodruff@maastrichtuniversity.nl (H.C.W.); a.ibrahim@maastrichtuniversity.nl (A.I.); philippe.lambin@maastrichtuniversity.nl (P.L.)

<sup>2</sup> Department of Diagnostic Radiology, Faculty of Applied Medical Sciences, Jazan University, Jazan 45142, Saudi Arabia

<sup>3</sup> Department of Radiology and Nuclear Medicine, Maastricht University Medical Center, 6200 Maastricht, The Netherlands; f.mottaghy@mumc.nl

<sup>4</sup> Division of Nuclear Medicine and Oncological Imaging, Department of Medical Physics, University Hospital of Liege and GIGA CRC-In Vivo Imaging, University of Liege, 4000 Liege, Belgium; rhustinx@chu.ulg.ac.be

<sup>5</sup> Department of Nuclear Medicine and Comprehensive Diagnostic Center Aachen (CDCA), University Hospital RWTH Aachen University, 52074 Aachen, Germany

<sup>6</sup> Department of Radiology, Columbia University Irving Medical Center, New York, NY 10032, USA

\* Correspondence: t.refae@maastrichtuniversity.nl

† These authors contributed equally to this work.

‡ These authors contributed equally to this work.



**Citation:** Refae, T.; Salahuddin, Z.; Widaatalla, Y.; Primakov, S.; Woodruff, H.C.; Hustinx, R.; Mottaghy, F.M.; Ibrahim, A.; Lambin, P. CT Reconstruction Kernels and the Effect of Pre- and Post-Processing on the Reproducibility of Handcrafted Radiomic Features. *J. Pers. Med.* **2022**, *12*, 553. <https://doi.org/10.3390/jpm12040553>

Academic Editors: Aristotelis Chatzioannou and Yudong Zhang

Received: 3 March 2022

Accepted: 28 March 2022

Published: 31 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Handcrafted radiomics features (HRFs) are quantitative features extracted from medical images to decode biological information to improve clinical decision making. Despite the potential of the field, limitations have been identified. The most important identified limitation, currently, is the sensitivity of HRF to variations in image acquisition and reconstruction parameters. In this study, we investigated the use of Reconstruction Kernel Normalization (RKN) and ComBat harmonization to improve the reproducibility of HRFs across scans acquired with different reconstruction kernels. A set of phantom scans ( $n = 28$ ) acquired on five different scanner models was analyzed. HRFs were extracted from the original scans, and scans were harmonized using the RKN method. ComBat harmonization was applied on both sets of HRFs. The reproducibility of HRFs was assessed using the concordance correlation coefficient. The difference in the number of reproducible HRFs in each scenario was assessed using McNemar's test. The majority of HRFs were found to be sensitive to variations in the reconstruction kernels, and only six HRFs were found to be robust with respect to variations in reconstruction kernels. The use of RKN resulted in a significant increment in the number of reproducible HRFs in 19 out of the 67 investigated scenarios (28.4%), while the ComBat technique resulted in a significant increment in 36 (53.7%) scenarios. The combination of methods resulted in a significant increment in 53 (79.1%) scenarios compared to the HRFs extracted from original images. Since the benefit of applying the harmonization methods depended on the data being harmonized, reproducibility analysis is recommended before performing radiomics analysis. For future radiomics studies incorporating images acquired with similar image acquisition and reconstruction parameters, except for the reconstruction kernels, we recommend the systematic use of the pre- and post-processing approaches (respectively, RKN and ComBat).

**Keywords:** radiomics reproducibility; reconstruction kernel; ComBat harmonization; image harmonization

## 1. Introduction

Recent decades have witnessed an exponentially increasing number of studies investigating the potential of quantitative imaging features to extract additional information from medical images not detectable by human eyes [1,2]. Handcrafted radiomics refers to the high-throughput extraction of quantitative imaging features from medical images to decode biologic information [3,4] and, today, more than 5000 studies can be returned on the PubMed database using “radiomics” as a search word. The handcrafted radiomics approach “involves manual segmentation of the region of interest (e.g., the tumor) on medical imaging and extraction of thousands of human-defined and curated quantitative features from the region of interest” [5].

The hypothesis in radiomics studies is that handcrafted radiomic features (HRFs) can be used singularly or collectively as clinical biomarkers [3]. Many studies have investigated and reported on the potential of HRFs to predict clinical endpoints, such as overall survival [6–8], tissue histology [9–13] and response to therapy [14,15]. These studies highlighted the potential of such approaches to be applied in clinical settings, since they could present non-invasive, reliable, readily available and cost-effective alternatives to current invasive clinical procedures, such as tissue biopsies. Moreover, with proper application, radiomics could provide reproducible predictions, which are quantitative and less dependent on the subjective interpretation of medical examinations [16,17].

With the development of handcrafted radiomics as a research field, the limitations the field faces have been increasingly investigated during recent years [4,18]. The most important identified limitation currently is the sensitivity of HRFs to variations in image acquisition and reconstruction parameters [19–24]. For an HRF to be used as a clinical biomarker (solely or in combination with other HRFs), it has to be reproducible across different imaging parameters for generalization purposes [24]. However, many studies have reported on the sensitivity of HRFs to variations in time (test–retest) [25–29] and to variations in imaging acquisition and reconstruction parameters [30–37]. Studies have also reported that the degree of variation in a single acquisition or reconstruction parameter affects the reproducibility of HRFs variably [31,34]. A number of studies have reported the significant effects of variations in reconstruction kernels on the reproducibility of HRFs [20,38].

Different methods have been investigated to address the issue of reproducibility of HRFs across scans acquired differently. ComBat harmonization [39] is one of the post-processing methods that have recently been extensively investigated in radiomics analyses [40–42]. ComBat harmonization is a method that was developed for removing batch effects—attributed to the use of different machinery—from gene expression arrays. A number of studies have reported on the applicability of ComBat harmonization in different scenarios, such as scans acquired with varying degrees of differences in CT image acquisition and reconstruction parameters, scans acquired with a single variation in an image reconstruction parameter (in-plane resolution) and scans of different contrast-enhancement phases [31,35,43,44]. These studies reported that the performance of ComBat in radiomics analyses is dependent on the variations in the data being harmonized. A number of studies have also investigated the potential of ComBat in different scenarios [45–48]. However, the potential of ComBat to remove batch effects attributed solely to the variations in the reconstruction kernel has yet to be thoroughly investigated. Other investigated methods include pre-processing of the images to minimize effects due to differences in slice thickness, reconstruction with convolutional kernels, etc. Normalization of chest CT data minimized the variability that resulted from different reconstruction kernels [49]. The authors developed a method that targeted reducing the variations in the quantification of emphysema by normalizing the reconstruction kernel (Reconstruction Kernel Normalization—RKN). The CT scans obtained from different scanners that were reconstructed with varying kernels showed reduced variability in emphysema quantification after the proposed iterative normalization. However, the effect of this normalization method on the reproducibility of HRFs has not been investigated.

In this study, we hypothesize that the use of RKN and ComBat could improve the reproducibility of HRFs across scans acquired with different reconstruction kernels depending on the variations in the data being analyzed and/or harmonized. We further hypothesize that the combination of both methods (RKN and ComBat) would give superior results in terms of “number of reproducible HRFs” compared to no or only one harmonization method. Given that variations in the convolution kernel impact the reproducibility of HRFs the most, we investigate the reproducibility of HRFs extracted from phantom CT scans acquired with different reconstruction kernels on different imaging vendors. We also investigate the potential of ComBat harmonization, RKN and the combination of both methods to reduce the variations in HRF values attributed to differences solely in the reconstruction kernels of the original scans.

## 2. Materials and Methods

### 2.1. Imaging Data

The phantom data used in the study were obtained from the public Credence Cartridge Radiomics (CCR) phantom dataset [50] from the Cancer Imaging Archive site (TCIA.org) [51]. A total of 251 scans were acquired using different scanners, acquisition and reconstruction parameters. For this study, we included scans that were acquired using the same imaging acquisition and reconstruction parameters, except for the convolution kernel. After applying the inclusion criteria, 28 scans from five different scanner models were used in this study (Table 1).

**Table 1.** Acquisition and reconstruction parameters for the imaging dataset.

Manufacturer	Scanner Model	Number of Scans	X-ray Tube Current (kV)	Convolution Kernels	Slice Thickness (mm)	Pixel Spacing (mm <sup>2</sup> )
GE	Discovery STE	5	120	Standard, Detail, Edge, Soft, Lung	1.25	0.49 × 0.49
Philips	Brilliance 64	4	120	A, B, C, L	1.50	0.49 × 0.49
Siemens	Sensation 40	6	120	B10f, B20f, B31f, B50f, B60f, B70f	1.50	0.49 × 0.49
	Sensation 64	7	120	B10f, B20f, B30f, B31f, B50f, B60f, B70f	1.50	0.49 × 0.49
	SOMATOM Definition AS	6	120	I26f, I30f, I40f, I44f, I50f, I70f	1.50	0.49 × 0.49

### 2.2. Volume of Interest and HRFs Extraction

Each layer of the phantom was segmented as a single volume of interest (VOI), with the dimensions  $8 \times 8 \times 2 \text{ cm}^3$ . A total of 10 VOIs were segmented per scan, resulting in a total of 280 VOIs. HRFs were extracted using the open source PyRadiomics software version 2.2.0 [52]. HRFs were extracted at two different stages: directly from the original scans; and after image pre-processing. Image intensities were binned in all of the three scenarios with a binwidth of 25 Hounsfield units (HUs) to reduce noise levels and texture matrix sizes and the amount of computational power needed. No other image pre-processing was applied in any of the scenarios. Extracted HRFs were HU intensity features and texture features of five matrices: gray-level co-occurrence (GLCM); gray-level run-length (GLRLM); gray-level size zone (GLSZM); gray-level dependence (GLDM); and neighborhood gray-tone difference (NGTDM) matrices. A more detailed description of PyRadiomics HRFs can be found online at: <https://pyradiomics.readthedocs.io/en/latest/features.html> (accessed on 13 October 2021).

### 2.3. Reconstruction Kernel Normalization

The CT scan  $I_0$  is decomposed into a series of frequency components  $F^i$ . Image  $I_0$  is convoluted with the Gaussian filter at  $\sigma_i$  scale ( $\sigma_i = 0, 1, 2, 4, 8, 16$ ) to get a filtered image  $L_{\sigma_i}$ . The frequency component for  $i = 0, 1, 2, 3, 4$  is given by  $F^{i+1} = L_{\sigma_{i+1}} - L_{\sigma_i}$  and for  $i = 5$  it is given by  $F^{i+1} = L_{\sigma_i}$ . The normalized image  $I_N$  is obtained by  $I_N = F^6 + \sum_{i=1}^5 \lambda_i \cdot F^5$ .  $\lambda_i$  is given by  $\frac{r_i}{e_i}$ , where  $r_i$  and  $e_i$  are the standard deviations of the intensity values in the band  $F^i$  of the reference image and image  $I_0$ , respectively. This process is repeated until  $\lambda_i$  is within the range  $[0.95, 1.05]$ . This method was proposed for reducing the effects of varying reconstruction kernels for emphysema quantification in chest CT scans [49]. We investigated the effect of applying this normalization method on feature reproducibility.

### 2.4. Image Pre-Processing and HRF Post-Processing

Four scenarios were analyzed in this study (Figure 1): (i) HRFs extracted from original images; (ii) HRFs extracted from pre-processed scans with the method described in Section 2.3; (iii) HRFs extracted from original images and harmonized with ComBat; and (iv) the combination of both methods. In scenario (ii), image pre-processing was performed using the method previously described in [49]. Each set of images ( $n = 5$ ) was normalized to a reference scan from the set. HRFs were extracted following image pre-processing. In scenario (iii), ComBat harmonization was applied on HRFs extracted from the original scans without pre-processing. ComBat harmonization in radiomics has been previously described [43]. In scenario (iv), HRFs were extracted from images normalized with the RKN method and harmonized using ComBat harmonization.

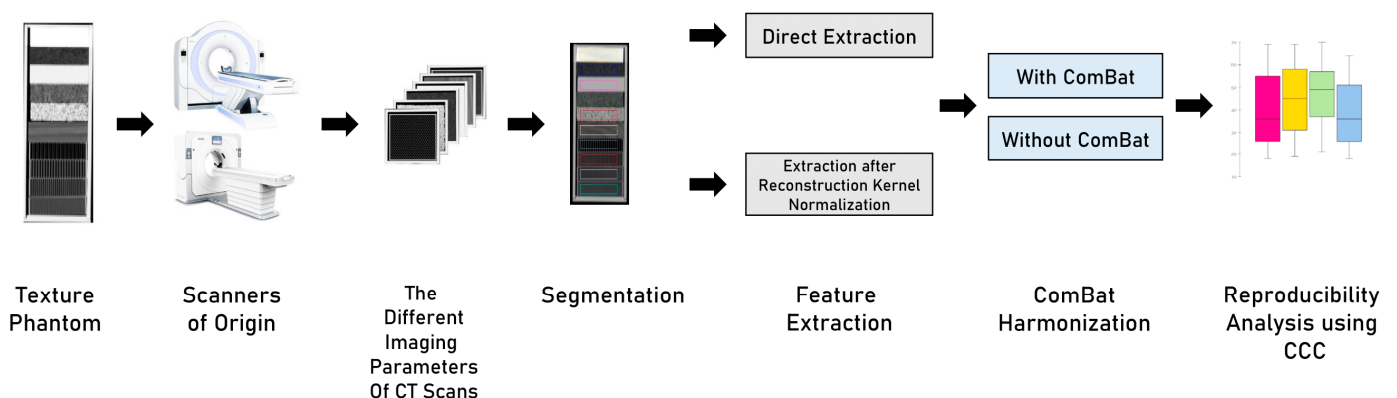


Figure 1. The study workflow.

### 2.5. Statistical Analysis

All statistical analyses were performed using R [53] on RStudio (V3.6.3) [54]. For each scanner model, scans were compared in a pair-wise manner. The concordance correlation coefficient (CCC) was used to assess the reproducibility of HRFs across different pairs [55] (epiR package V. 2.0.26) [56]. The CCC assesses the agreement in the value and rank for each HRF across the pairwise scenarios. HRFs with  $CCC > 0.9$  were considered reproducible in a given scenario. The CCC was calculated in each of the investigated scenarios described in Section 2.4.

To assess the statistical significance of the differences in the number of reproducible HRFs in each scenario, the McNemar test was used [57]. The McNemar test is used to assess whether marginal frequencies are equal before and after an intervention. In this study, we calculated McNemar’s  $p$ -values using the HRFs extracted from the original images and after RKN, ComBat, and the combination of both. We also calculated the  $p$ -values among the methods, as well as the  $p$ -values for each method compared to the combination of

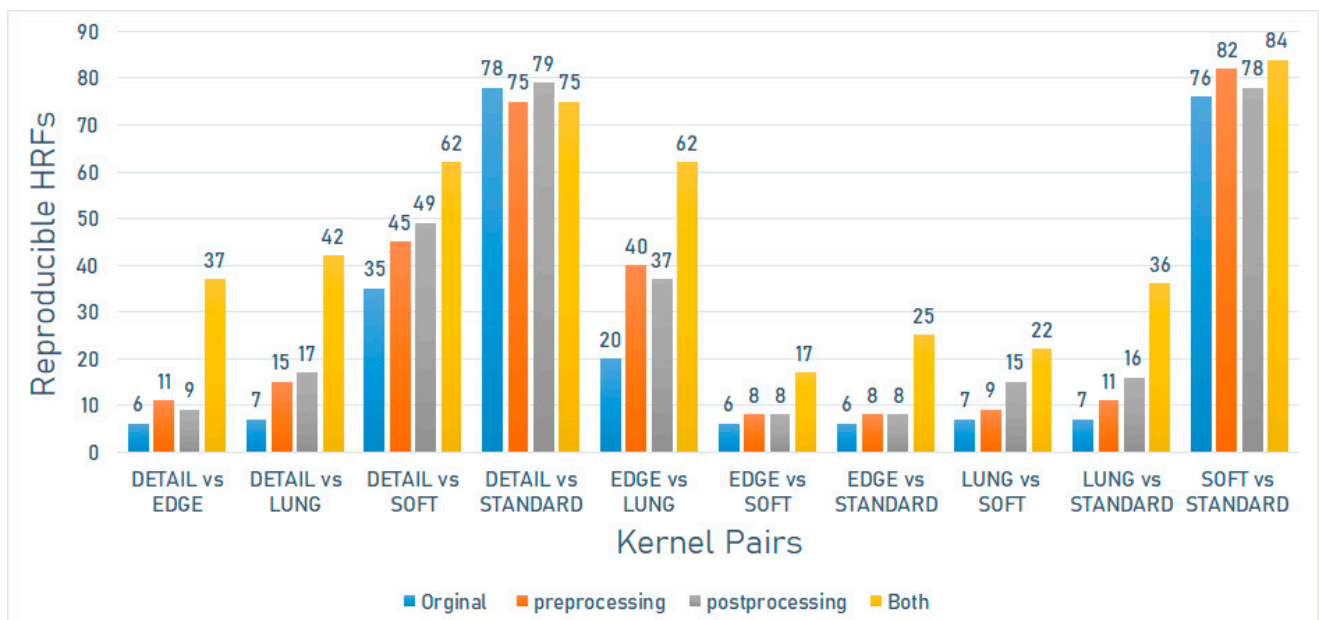
methods. For each pair, the difference in the number of reproducible HRFs was labeled “significant” or “not significant” depending on the *p*-value.

### 3. Results

#### 3.1. The Effect of Differences in Convolution Kernels on the Reproducibility of HRFs

The Pyradiomics toolbox provides a set of 91 original HRFs from each VOI. These HRFs are divided into First Order Statistics (*n* = 18), GLCM (*n* = 22), GLRLM (*n* = 16); GLSZM (*n* = 16), NGTDM (*n* = 5) and GLDM (*n* = 14). The number of reproducible HRFs varied across kernels and scanner models. Six HRFs were found to be robust to changes in convolution kernels across all scanner models: “Firstorder\_10Percentile”, “Firstorder\_Energy”, “Firstorder\_Mean”, “Firstorder\_Median”, “Firstorder\_RootMeanSquared” and “Firstorder\_TotalEnergy”.

On the Discovery STE scanner model (GE Medical Systems), the number of reproducible HRFs varied between 6 (6.59%) and 78 (85.71%). The greatest number of reproducible HRFs was observed across scans acquired with Detailed and Standard kernels (Figure 2).



**Figure 2.** The number of reproducible HRFs across different kernels on the Discovery STE scanner model.

On the Sensation 40 scanner model (Siemens), the number of reproducible HRFs varied between 6 (6.59%) and 91 (100%). The greatest number of reproducible HRFs was observed across scans acquired with B60f and B70f kernels (Figure 3).

On the SOMATOM definition scanner model (Siemens), the number of reproducible HRFs varied between 6 (6.59%) and 65 (71.4%). The greatest number of reproducible HRFs was observed across scans acquired with I44f and I50f kernels (Figure 4).



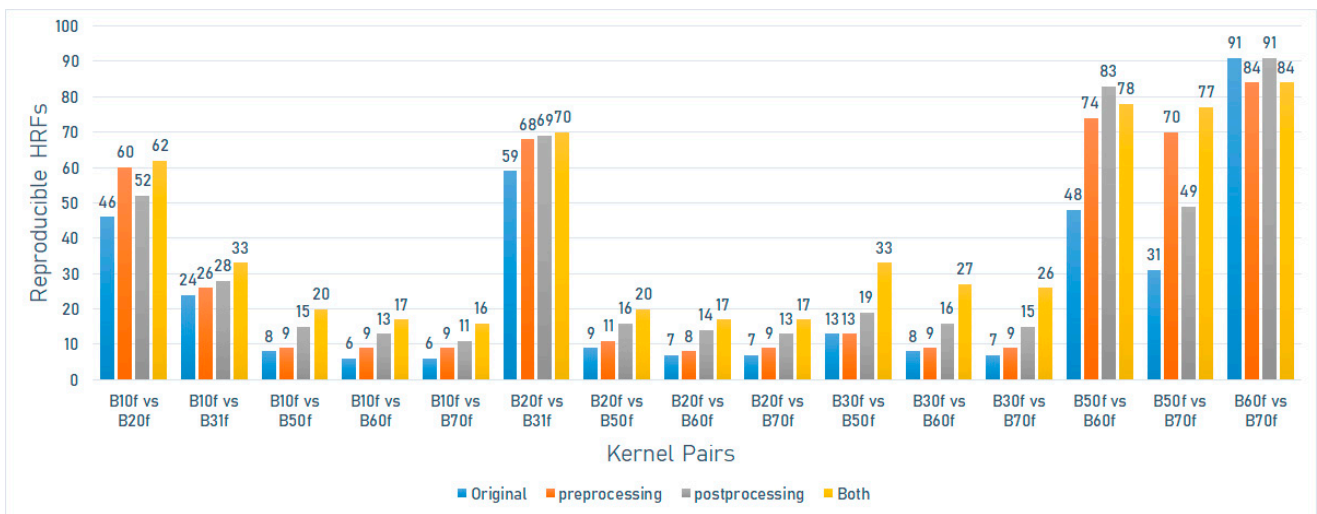


Figure 3. The number of reproducible HRFs across different kernels on the Sensation 40 scanner model.

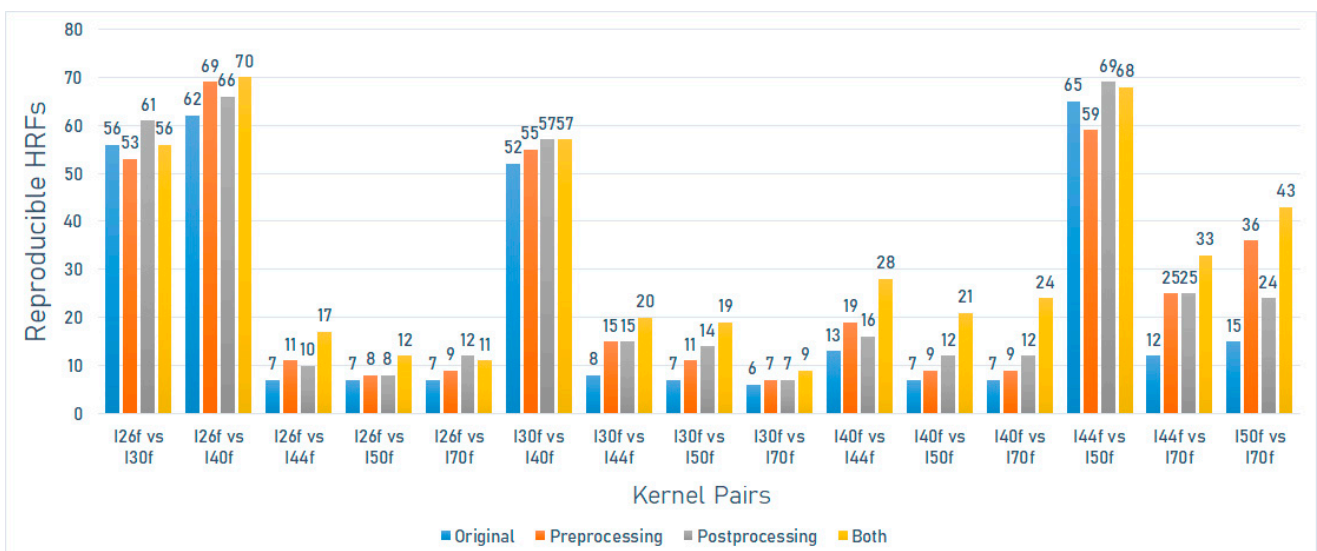


Figure 4. The number of reproducible HRFs across different kernels on the SOMATOM Definition scanner model.

On the Sensation 64 scanner model (Siemens), the number of reproducible HRFs varied between 6 (6.59%) and 91 (100%). The greatest number of reproducible HRFs was observed across scans acquired with B60f and B70f kernels (Figure 5).

On the Brilliance 64 scanner model (Philips), the number of reproducible HRFs varied between 14 (15.4%) and 48 (52.7%). The greatest number of reproducible HRFs was observed across scans acquired with A and B kernels (Figure 6).

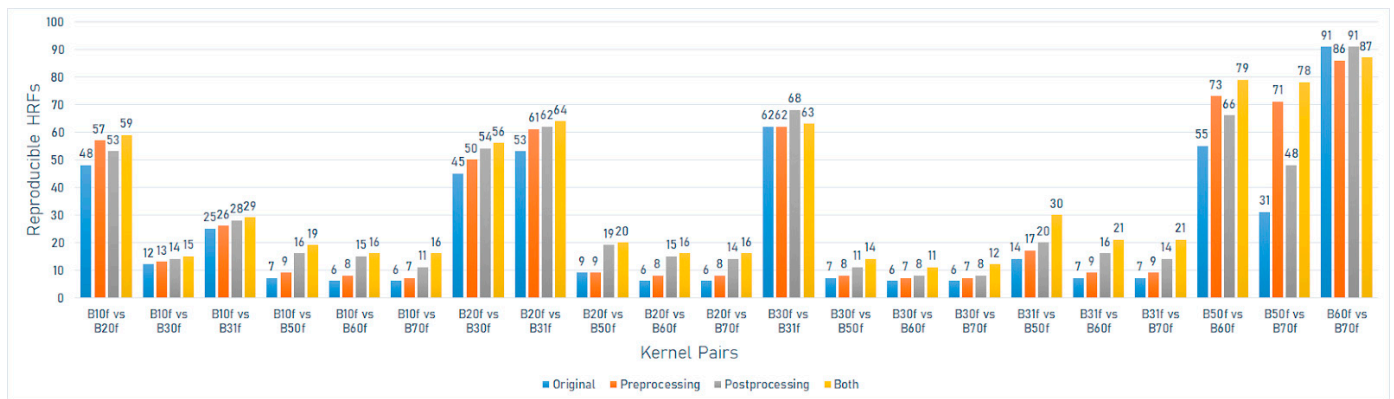


Figure 5. The number of reproducible HRFs across different kernels on the Sensation 64 scanner model.

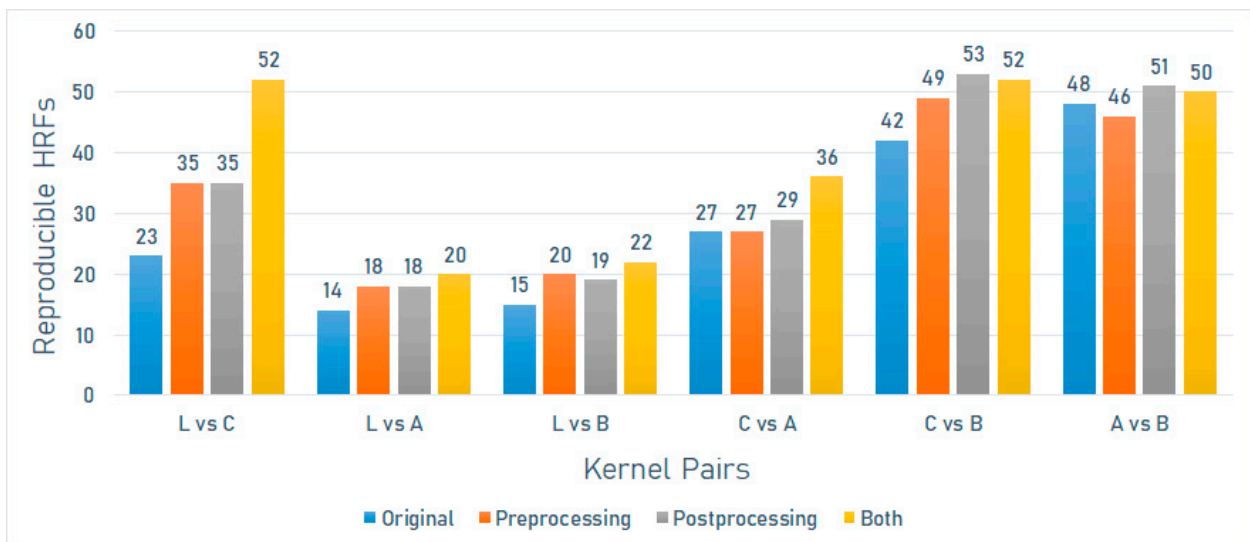


Figure 6. The number of reproducible HRFs across different kernels on the Brilliance 64 scanner model.

### 3.2. The Effects of Pre- and Post-Processing

#### 3.2.1. Reconstruction Kernel Normalization (RKN)

The number of HRFs that became reproducible following the application of the described method varied with the variations in kernels being harmonized and the scanner model used. In most of the investigated scenarios (58 out of 67; 86.6%), the use of this method has resulted in an increment in the number of reproducible HRFs. However, only 19 scenarios (28.4%) showed statistically significant increments. In a number of scenarios (6 out of the analyzed 67 scenarios (9%)), there was a net loss in the number of reproducible HRFs compared to the original, 2 (3%) of which were statistically significant (Figures 2–6). In three (4.5%) scenarios, there was no difference between the number of reproducible HRFs extracted from the original and the normalized images.

On the Discovery STE scanner model (GE Medical Systems), the number of reproducible HRFs extracted from the scans after image pre-processing varied between 8 (8.8%) and 82 (90.1%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with Edge and Lung kernels (Figure 2).

On the Sensation 40 scanner model (Siemens), the number of reproducible HRFs extracted from the scans after image pre-processing varied between 8 (8.8%) and 84 (92.3%).

In this scenario, the highest number of reproducible HRFs decreased compared to those extracted from the original images for the scans acquired with B60f and B70f. The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with B50f and B70f kernels (Figure 3).

On the SOMATOM definition scanner model (Siemens), the number of reproducible HRFs extracted from the scans after image pre-processing varied between 7 (7.7%) and 69 (75.8%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with I50f and I70f kernels (Figure 4).

On the Sensation 64 scanner model (Siemens), the number of reproducible HRFs extracted from the scans after image pre-processing varied between 7 (7.7%) and 86 (94.5%). In this scenario, the highest number of reproducible HRFs decreased compared to those extracted from the original images (B60f vs. B70f) (Figure 5).

On the Brilliance 64 scanner model (Philips), the number of reproducible HRFs extracted from the scans after image pre-processing varied between 18 (19.8%) and 49 (53.8%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with L and C kernels (Figure 6).

### 3.2.2. ComBat Harmonization

In 65 out of the 67 investigated scenarios (97%), there was a net increase in the number of reproducible HRFs compared to the original, with 36 (53.7%) scenarios witnessing significant statistical increments. In two scenarios, the same number of reproducible HRFs was found before and after ComBat harmonization. In 46 (68.7%) scenarios, ComBat harmonization outperformed the RKN method, 17 (25.4%) of which were statistically significant. In 13 (19.4%) scenarios, the RKN method outperformed ComBat harmonization, 5 (7.5%) of which were statistically significant increments.

On the Discovery STE scanner model (GE Medical Systems), the number of reproducible HRFs extracted from the scans after ComBat harmonization varied between 9 (9.9%) and 79 (86.8%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with Edge and Lung kernels (Figure 2).

On the Sensation 40 scanner model (Siemens), the number of reproducible HRFs extracted from the scans after ComBat harmonization varied between 11 (12.1%) and 69 (75.8%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with B50f and B60f kernels (Figure 3).

On the SOMATOM definition scanner model (Siemens), the number of reproducible HRFs extracted from the scans after ComBat harmonization pre-processing varied between 7 (7.7%) and 69 (75.8%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with I44f and I70f kernels (Figure 4).

On the Sensation 64 scanner model (Siemens), the number of reproducible HRFs extracted from the scans after ComBat harmonization varied between 8 (8.8%) and 91 (100%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with B50f and B70f kernels (Figure 5).

On the Brilliance 64 scanner model (Philips), the number of reproducible HRFs extracted from the scans after ComBat harmonization varied between 18 (19.8%) and 53 (58.8%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with L and C kernels (Figure 6).

### 3.2.3. The Combination of Pre- and Post-Processing

In 63 (95.5%) out of the 67 investigated scenarios, there was a net increase in the number of reproducible HRFs compared to the original, 53 (79.1%) of which were statistically significant. Three (4.5%) showed a lower number of reproducible HRFs, with one (1.5%) scenario showing significantly fewer ( $p < 0.05$ ). The same number of reproducible HRFs was observed in one (1.5%) scenario. In 66 (98.5%) scenarios, the combination of methods



outperformed the RKN method, with 42 (62.7%) being significantly higher. The same number of reproducible HRFs was observed in one (1.5%) scenario. With regards to ComBat harmonization, the combination of methods resulted in a higher number of reproducible HRFs in 56 (83.6%) scenarios, 27 (40.3%) of which were statistically significant. A higher number of reproducible HRFs was obtained using only ComBat harmonization in 10 (14.9%) scenarios, only one (1.5%) of which was statistically significant. The same number of reproducible HRFs was observed in one (1.5%) scenario.

On the Discovery STE scanner model (GE Medical Systems), the number of reproducible HRFs extracted from the normalized scans after ComBat harmonization varied between 17 (18.7%) and 84 (92.3%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with Edge and Lung kernels (Figure 2).

On the Sensation 40 scanner model (Siemens), the number of reproducible HRFs extracted from the normalized scans after ComBat harmonization varied between 16 (17.6%) and 84 (92.3%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with B50f and B70f kernels (Figure 3).

On the SOMATOM definition scanner model (Siemens), the number of reproducible HRFs extracted from the normalized scans after ComBat harmonization pre-processing varied between 9 (9.9%) and 70 (77%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with I50f and I70f kernels (Figure 4).

On the Sensation 64 scanner model (Siemens), the number of reproducible HRFs extracted from the normalized scans after ComBat harmonization varied between 11 (12.1%) and 87 (95.7%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with B50f and B70f kernels (Figure 5).

On the Brilliance 64 scanner model (Philips), the number of reproducible HRFs extracted from the normalized scans after ComBat harmonization varied between 20 (22%) and 52 (57.2%). The greatest increment in the number of reproducible HRFs compared to the original images was observed across scans acquired with L and C kernels (Figure 6).

#### 4. Discussion

In this study, we analyzed the effects of difference in convolution kernels on five different scanner models, when all other CT acquisition and reconstruction parameters were fixed on a phantom dataset. We further investigated the ability of an image pre-processing (iterative normalization by frequency decomposition) method, and an HRF post-processing harmonization (using ComBat harmonization) method. Our results showed significant differences in the number of reproducible HRFs across the investigated scenarios. Scans reconstructed with similar convolution kernels showed a higher number of reproducible HRFs compared to scans reconstructed with significantly different convolution kernels. Similarly, the performance of both harmonization methods investigated varied with the differences in convolution kernels of the scans being harmonized.

Siemens scanner models (Sensation 40 and 64) have shown the reproducibility of all HRFs across the scans acquired with the higher end of convolution kernels (B60 and B70). Convolution kernels at the opposite end of the spectrum (for example, B10 and B70 on Siemens scanners) have shown the lowest number of reproducible HRFs. As such, our results are in line with previous studies that reported that the reproducibility of HRFs can be significantly affected by variations in convolution kernels [38,58–60].

The use of the RKN method on our dataset has resulted in a range of effects on the number of reproducible features, from negative to neutral to positive, depending on the scans being compared. We have observed a significant increase in the number of reproducible HRFs in most scenarios and a decrease in the number of reproducible HRFs in some other scenarios. This could be justified by the possibility that the analyzed data

in this study included a wider range of convolution kernels than those used to develop the method.

The application of ComBat harmonization resulted in a higher number of reproducible HRFs compared to those before harmonization in almost all of the investigated scenarios, which is in line with previous reports [43,44,61]. Moreover, on average, ComBat harmonization outperformed the image pre-processing method. The performance of ComBat further depended on the differences in the convolution kernels of the scans being harmonized. In general, the number of reproducible HRFs after ComBat harmonization followed a similar pattern to that of the number of reproducible HRFs before post-processing. These findings are in line with previous studies that investigated the applicability of ComBat harmonization in radiomics analyses [31,34]. The results add to the evidence on the need for reproducibility analyses in radiomics studies, including scans acquired differently, as well as the need for radiomics-specific harmonization methods.

The combination of RKN and ComBat harmonization methods resulted in a higher number of reproducible HRFs across the majority of the investigated scenarios. This indicates that each method could be addressing the reproducibility of HRFs in different manners, with their having been shown to be complementary to each other in many of the investigated scenarios. Nevertheless, the combination resulted in a lower number of reproducible HRFs in an appreciated percentage of scenarios compared to ComBat harmonization only. This suggests the need for reproducibility analysis before applying harmonization methods in radiomics analyses.

We identified six HRFs that were robust with respect to variations in convolution kernels across all the investigated scenarios. These HRFs were first-order statistics, and their robustness could be justified by the standardization of HUs across scanners. However, the majority of texture HRFs were sensitive to the majority of variations in convolution kernels. Clear to the eye, the standardization of image acquisition and reconstruction parameters would be the cornerstone for the translation of radiomic signatures to clinical practice. The findings of this study, and previous experiments, have shown that the reproducibility of HRFs significantly depends on imaging acquisition and reconstruction parameters. Therefore, reproducibility analysis is needed for a proper understanding of their performance or generalizability [19]. Another potential solution would be the development of radiomic signatures specific to a set of imaging acquisition and reconstruction parameters. However, this solution limits the generalizability of radiomic signatures.

While we tried to analyze all the kernels used in clinical practice, we were limited by the available data. However, the results have shown a similar pattern across different scanner models. Future studies that include a wider spectrum of convolution kernels are recommended. Furthermore, we limited our analyses to the original HRFs as they are commonly standardized across radiomics platforms. Detailed full HRF reproducibility analysis could be beneficial for specific tasks. Furthermore, the analysis was performed on a phantom dataset that was designed to mimic human tissues. However, it only gives an idea about the reproducibility of HRFs in the given scenarios, and similar analysis is needed for patient datasets to gain a full understanding. The potential of other harmonization methods, for example, dynamic range limitation [62], could also be explored in future studies. Additionally, the sensitivity of HRFs to variations in segmentations could not be assessed in this study, due to the use of automated segmentations.

## 5. Conclusions

The reproducibility of the majority of HRFs depended on the variations in reconstruction kernels in the data being analyzed. Six HRFs were found to be reproducible across all investigated scenarios. Radiomics analysis of scans acquired with different reconstruction kernels is not recommended in the absence of reproducibility analysis. We recommend the systematic use of RKN and ComBat harmonization in future radiomics studies, including images acquired similarly except for the reconstruction kernel. Nevertheless, their application should follow a reproducibility analysis to identify the set of reproducible HRFs

after harmonization. HRF-specific harmonization methods remain necessities in the field of radiomics.

**Author Contributions:** Conceptualization, T.R., A.I. and P.L.; methodology, T.R., A.I. and P.L.; software, T.R., Z.S., Y.W. and S.P.; formal analysis, T.R. and Z.S.; data curation, T.R. and Y.W.; writing—original draft preparation, T.R., Z.S. and Y.W.; writing—review and editing, S.P., H.C.W., R.H., F.M.M., A.I. and P.L.; visualization, T.R., Z.S., Y.W. and S.P.; supervision: A.I. and P.L.; project administration, A.I. and P.L.; funding acquisition, P.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors acknowledge financial support from ERC advanced grant (ERC-ADG-2015 n° 694812—Hypoximmuno), ERC-2020-PoC: 957565-AUTO.DISTINCT. The authors also acknowledge financial support from the European Union’s Horizon 2020 research and innovation program under grant agreement: ImmunoSABR n° 733008, MSCA-ITN-PREDICT n° 766276, CHAIMELEON n° 952172, EuCanImage n° 952103, JTI-IMI2-2020-23-two-stage IMI-OPTIMA n° 101034347 and TRANSCAN Joint Transnational Call 2016 (JTC2016 CLEARLY n° UM 2017-8295).

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are openly available at TCIA.org: <http://doi.org/10.7937/K9/TCIA.2017.zuzrml5b> (accessed on 7 January 2021).

**Conflicts of Interest:** Philippe Lambin reports, within and outside the submitted work, grants/sponsored research agreements from Radiomics SA, ptTheragnostic/DNAmito and Health Innovation Ventures. He received an advisor/presenter fee and/or reimbursement of travel costs/consultancy fee and/or in-kind manpower contribution from Radiomics SA, BHV, Merck, Varian, Elekta, ptTheragnostic, BMS and Convert pharmaceuticals. Lambin has minority shares in the companies Radiomics SA, Convert pharmaceuticals, Comunicare and LivingMed Biotech, and he is co-inventor of two issued patents with royalties on radiomics (PCT/NL2014/050248, PCT/NL2014/050728), licensed to Radiomics SA, and one issued patent on mtDNA (PCT/EP2014/059089), licensed to ptTheragnostic/DNAmito, three non-patented inventions (softwares) licensed to ptTheragnostic/DNAmito, Radiomics SA and Health Innovation Ventures and three non-issued, non-licensed patents on Deep Learning-Radiomics and LSRT (N2024482, N2024889, N2024889). He confirms that none of the above entities or funding sources were involved in the preparation of this paper. Woodruff has minority shares in the company Radiomics SA. The rest of the co-authors declare no competing interest.

## References

- Walsh, S.; de Jong, E.E.C.; van Timmeren, J.E.; Ibrahim, A.; Compter, I.; Peerlings, J.; Sanduleanu, S.; Refaee, T.; Keek, S.; Larue, R.T.H.M.; et al. Decision support systems in oncology. *JCO Clin. Cancer Inform.* **2019**, *3*, 1–9. [[CrossRef](#)] [[PubMed](#)]
- Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [[CrossRef](#)]
- Lambin, P.; Rios-Velazquez, E.; Leijenaar, R.; Carvalho, S.; van Stiphout, R.G.P.M.; Granton, P.; Zegers, C.M.L.; Gillies, R.; Boellard, R.; Dekker, A.; et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **2012**, *48*, 441–446. [[CrossRef](#)]
- Yip, S.S.F.; Aerts, H.J.W.L. Applications and limitations of radiomics. *Phys. Med. Biol.* **2016**, *61*, R150–R166. [[CrossRef](#)] [[PubMed](#)]
- Hosny, A.; Aerts, H.J.; Mak, R.H. Handcrafted versus deep learning radiomics for prediction of cancer therapy response. *Lancet Digit. Health* **2019**, *1*, e106–e107. [[CrossRef](#)]
- Aerts, H.J.W.L.; Velazquez, E.R.; Leijenaar, R.T.H.; Parmar, C.; Grossmann, P.; Carvalho, S.; Bussink, J.; Monshouwer, R.; Haibe-Kains, B.; Rietveld, D.; et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **2014**, *5*, 4006. [[CrossRef](#)]
- Bae, S.; Choi, Y.S.; Ahn, S.S.; Chang, J.H.; Kang, S.-G.; Kim, E.H.; Kim, S.H.; Lee, S.-K. Radiomic MRI phenotyping of glioblastoma: Improving survival prediction. *Radiology* **2018**, *289*, 797–806. [[CrossRef](#)] [[PubMed](#)]
- Oikonomou, A.; Khalvati, F.; Tyrrell, P.N.; Haider, M.A.; Tarique, U.; Jimenez-Juan, L.; Tjong, M.C.; Poon, I.; Eilaghi, A.; Ehrlich, L.; et al. Radiomics analysis at PET/CT contributes to prognosis of recurrence and survival in lung cancer treated with stereotactic body radiotherapy. *Sci. Rep.* **2018**, *8*, 4003. [[CrossRef](#)]
- Wu, W.; Parmar, C.; Grossmann, P.; Quackenbush, J.; Lambin, P.; Bussink, J.; Mak, R.; Aerts, H.J.W.L. Exploratory study to identify radiomics classifiers for lung cancer histology. *Front. Oncol.* **2016**, *6*, 71. [[CrossRef](#)] [[PubMed](#)]
- Blüthgen, C.; Patella, M.; Euler, A.; Baessler, B.; Martini, K.; von Spiczak, J.; Schneider, D.; Opitz, I.; Frauenfelder, T. Computed tomography radiomics for the prediction of thymic epithelial tumor histology, TNM stage and myasthenia gravis. *PLoS ONE* **2021**, *16*, e0261401. [[CrossRef](#)]

11. Linning, E.; Lu, L.; Li, L.; Yang, H.; Schwartz, L.H.; Zhao, B. Radiomics for Classification of Lung Cancer Histological Subtypes Based on Nonenhanced Computed Tomography. *Acad. Radiol.* **2019**, *26*, 1245–1252. [[CrossRef](#)]
12. Stefan, P.-A.; Puscas, M.E.; Csuk, C.; Lebovici, A.; Petresc, B.; Lupean, R.; Mihiu, C.M. The utility of texture-based classification of different types of ascites on magnetic resonance. *J. BUON* **2020**, *25*, 1237–1244. [[PubMed](#)]
13. Csutak, C.; Ștefan, P.-A.; Lupean, R.-A.; Lenghel, L.M.; Mihiu, C.M.; Lebovici, A. Computed tomography in the diagnosis of intraperitoneal effusions: The role of texture analysis. *Bosn. J. Basic Med. Sci.* **2021**, *21*, 488–494. [[CrossRef](#)] [[PubMed](#)]
14. Horvat, N.; Veeraraghavan, H.; Khan, M.; Blazic, I.; Zheng, J.; Capanu, M.; Sala, E.; Garcia-Aguilar, J.; Gollub, M.J.; Petkovska, I. MR imaging of rectal cancer: Radiomics analysis to assess treatment response after neoadjuvant therapy. *Radiology* **2018**, *287*, 833–843. [[CrossRef](#)] [[PubMed](#)]
15. Tharmalingam, H.; Tsang, Y.M.; Alonzi, R.; Beasley, W.; Taylor, N.J.; McWilliam, A.; Padhani, A.; Choudhury, A.; Hoskin, P.J. Changes in magnetic resonance imaging radiomic features in response to androgen deprivation therapy in patients with intermediate- and high-risk prostate cancer. *Clin. Oncol.* **2022**. [[CrossRef](#)] [[PubMed](#)]
16. Ștefan, P.-A.; Lupean, R.-A.; Mihiu, C.M.; Lebovici, A.; Oancea, M.D.; Hițu, L.; Duma, D.; Csutak, C. Ultrasonography in the diagnosis of adnexal lesions: The role of texture analysis. *Diagnostics* **2021**, *11*, 812. [[CrossRef](#)] [[PubMed](#)]
17. Ștefan, R.-A.; Ștefan, P.-A.; Mihiu, C.M.; Csutak, C.; Melincovici, C.S.; Crivii, C.B.; Maluțan, A.M.; Hițu, L.; Lebovici, A. Ultrasonography in the differentiation of endometriomas from hemorrhagic ovarian cysts: The role of texture analysis. *J. Pers. Med.* **2021**, *11*, 611. [[CrossRef](#)]
18. Lohmann, P.; Bousabarah, K.; Hoevens, M.; Treuer, H. Radiomics in radiation oncology—Basics, methods, and limitations. *Strahlenther. Onkol.* **2020**, *196*, 848–855. [[CrossRef](#)] [[PubMed](#)]
19. Ibrahim, A.; Primakov, S.; Beuque, M.; Woodruff, H.C.; Halilaj, I.; Wu, G.; Refaee, T.; Granzier, R.; Widaatalla, Y.; Hustinx, R.; et al. Radiomics for precision medicine: Current challenges, future prospects, and the proposal of a new framework. *Methods* **2020**, *188*, 20–29. [[CrossRef](#)]
20. Mali, S.A.; Ibrahim, A.; Woodruff, H.C.; Andrearczyk, V.; Müller, H.; Primakov, S.; Salahuddin, Z.; Chatterjee, A.; Lambin, P. Making radiomics more reproducible across scanner and imaging protocol variations: A review of harmonization methods. *J. Pers. Med.* **2021**, *11*, 842. [[CrossRef](#)] [[PubMed](#)]
21. Midya, A.; Chakraborty, J.; Gönen, M.; Do, R.K.G.; Simpson, A.L. Influence of CT acquisition and reconstruction parameters on radiomic feature reproducibility. *J. Med. Imaging* **2018**, *5*, 011020. [[CrossRef](#)] [[PubMed](#)]
22. Reiazi, R.; Abbas, E.; Famiyeh, P.; Rezaie, A.; Kwan, J.Y.Y.; Patel, T.; Bratman, S.V.; Tadic, T.; Liu, F.-F.; Haibe-Kains, B. The impact of the variation of imaging parameters on the robustness of computed tomography radiomic features: A review. *Comput. Biol. Med.* **2021**, *133*, 104400. [[CrossRef](#)] [[PubMed](#)]
23. Espinasse, M.; Pitre-Champagnat, S.; Charmettant, B.; Bidault, F.; Volk, A.; Balleyguier, C.; Lassau, N.; Caramella, C. CT texture analysis challenges: Influence of acquisition and reconstruction parameters: A comprehensive review. *Diagnostics* **2020**, *10*, 258. [[CrossRef](#)] [[PubMed](#)]
24. Zhao, B. Understanding sources of variation to improve the reproducibility of radiomics. *Front. Oncol.* **2021**, *11*, 633176. [[CrossRef](#)] [[PubMed](#)]
25. Granzier, R.W.Y.; Ibrahim, A.; Primakov, S.; Keek, S.A.; Halilaj, I.; Zwanenburg, A.; Engelen, S.M.E.; Lobbes, M.B.I.; Lambin, P.; Woodruff, H.C.; et al. Test-retest data for the assessment of breast MRI radiomic feature repeatability. *J. Magn. Reson. Imaging* **2021**. [[CrossRef](#)]
26. Shiri, I.; Abdollahi, H.; Shaysteh, S.; Mahdavi, S.R. Test-retest reproducibility and robustness analysis of recurrent glioblastoma MRI radiomics texture features. *Iran. J. Radiol.* **2017**, *5*, e48035. [[CrossRef](#)]
27. Peerlings, J.; Woodruff, H.C.; Winfield, J.M.; Ibrahim, A.; van Beers, B.E.; Heerschap, A.; Jackson, A.; Wildberger, J.E.; Mottaghy, F.M.; DeSouza, N.M.; et al. Stability of radiomics features in apparent diffusion coefficient maps from a multi-centre test-retest trial. *Sci. Rep.* **2019**, *9*, 4800. [[CrossRef](#)] [[PubMed](#)]
28. Pfaehler, E.; Beukinga, R.J.; de Jong, J.R.; Slart, R.H.J.A.; Slump, C.H.; Dierckx, R.A.J.O.; Boellaard, R. Repeatability of 18 F-FDG PET radiomic features: A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method. *Med. Phys.* **2019**, *46*, 665–678. [[CrossRef](#)]
29. Prayer, F.; Hofmanninger, J.; Weber, M.; Kifjak, D.; Willenpart, A.; Pan, J.; Röhrich, S.; Langs, G.; Prosch, H. Variability of computed tomography radiomics features of fibrosing interstitial lung disease: A test-retest study. *Methods* **2021**, *188*, 98–104. [[CrossRef](#)] [[PubMed](#)]
30. Zhao, B.; Tan, Y.; Tsai, W.-Y.; Qi, J.; Xie, C.; Lu, L.; Schwartz, L.H. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci. Rep.* **2016**, *6*, 23428. [[CrossRef](#)] [[PubMed](#)]
31. Ibrahim, A.; Refaee, T.; Primakov, S.; Barufaldi, B.; Acciavatti, R.J.; Granzier, R.W.Y.; Hustinx, R.; Mottaghy, F.M.; Woodruff, H.C.; Wildberger, J.E.; et al. The effects of in-plane spatial resolution on CT-based radiomic features' stability with and without ComBat harmonization. *Cancers* **2021**, *13*, 1848. [[CrossRef](#)] [[PubMed](#)]
32. Zhovannik, I.; Bussink, J.; Traverso, A.; Shi, Z.; Kalendralis, P.; Wee, L.; Dekker, A.; Fijten, R.; Monshouwer, R. Learning from scanners: Bias reduction and feature correction in radiomics. *Clin. Transl. Radiat. Oncol.* **2019**, *19*, 33–38. [[CrossRef](#)] [[PubMed](#)]



33. Ibrahim, A.; Primakov, S.; Barufaldi, B.; Acciavatti, R.J.; Granzier, R.W.Y.; Hustinx, R.; Mottaghy, F.M.; Woodruff, H.C.; Wildberger, J.E.; Lambin, P.; et al. Reply to Orlhac, F.; Buvat, I. Comment on “Ibrahim et al. The Effects of In-Plane Spatial Resolution on CT-Based Radiomic Features’ Stability with and without ComBat Harmonization. *Cancers* **2021**, *13*, 1848”. *Cancers* **2021**, *13*, 3080. [[CrossRef](#)] [[PubMed](#)]
34. Ibrahim, A.; Refaee, T.; Leijenaar, R.T.H.; Primakov, S.; Hustinx, R.; Mottaghy, F.M.; Woodruff, H.C.; Maidment, A.D.A.; Lambin, P. The application of a workflow integrating the variable reproducibility and harmonizability of radiomic features on a phantom dataset. *PLoS ONE* **2021**, *16*, e0251147. [[CrossRef](#)]
35. Ibrahim, A.; Widaatalla, Y.; Refaee, T.; Primakov, S.; Miclea, R.L.; Öcal, O.; Fabritius, M.P.; Ingrisich, M.; Rieke, J.; Hustinx, R.; et al. Reproducibility of CT-based hepatocellular carcinoma radiomic features across different contrast imaging phases: A proof of concept on SORAMIC trial data. *Cancers* **2021**, *13*, 4638. [[CrossRef](#)]
36. Park, J.E.; Park, S.Y.; Kim, H.J.; Kim, H.S. Reproducibility and generalizability in radiomics modeling: Possible strategies in radiologic and statistical perspectives. *Korean J. Radiol.* **2019**, *20*, 1124–1137. [[CrossRef](#)] [[PubMed](#)]
37. Meyer, M.; Ronald, J.; Vernuccio, F.; Nelson, R.C.; Ramirez-Giraldo, J.C.; Solomon, J.; Patel, B.N.; Samei, E.; Marin, D. Reproducibility of CT radiomic features within the same patient: Influence of radiation dose and CT reconstruction settings. *Radiology* **2019**, *293*, 583–591. [[CrossRef](#)] [[PubMed](#)]
38. Lu, L.; Ehmke, R.C.; Schwartz, L.H.; Zhao, B. Assessing agreement between radiomic features computed for multiple CT imaging settings. *PLoS ONE* **2016**, *11*, e0166550. [[CrossRef](#)]
39. Johnson, W.E.; Li, C.; Rabinovic, A. Adjusting batch effects in microarray expression data using empirical bayes methods. *Biostatistics* **2007**, *8*, 118–127. [[CrossRef](#)]
40. Ligeró, M.; Jordi-Ollero, O.; Bernatowicz, K.; Garcia-Ruiz, A.; Delgado-Muñoz, E.; Leiva, D.; Mast, R.; Suarez, C.; Sala-Llonch, R.; Calvo, N.; et al. Minimizing acquisition-related radiomics variability by image resampling and batch effect correction to allow for large-scale data analysis. *Eur. Radiol.* **2021**, *31*, 1460–1470. [[CrossRef](#)]
41. Foy, J.J.; Al-Hallaq, H.A.; Grekoski, V.; Tran, T.; Guruvadoo, K.; Armato, S.G., III; Sensakovic, W.F. Harmonization of radiomic feature variability resulting from differences in CT image acquisition and reconstruction: Assessment in a cadaveric liver. *Phys. Med. Biol.* **2020**, *65*, 205008. [[CrossRef](#)] [[PubMed](#)]
42. Arendt, C.T.; Leithner, D.; Mayerhoefer, M.E.; Gibbs, P.; Czerny, C.; Arnoldner, C.; Burck, I.; Leinung, M.; Tanyildizi, Y.; Lenga, L.; et al. Radiomics of high-resolution computed tomography for the differentiation between cholesteatoma and middle ear inflammation: Effects of post-reconstruction methods in a dual-center study. *Eur. Radiol.* **2021**, *31*, 4071–4078. [[CrossRef](#)]
43. Fortin, J.-P.; Parker, D.; Tunç, B.; Watanabe, T.; Elliott, M.A.; Ruparel, K.; Roalf, D.R.; Satterthwaite, T.D.; Gur, R.C.; Gur, R.E.; et al. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* **2017**, *161*, 149–170. [[CrossRef](#)] [[PubMed](#)]
44. Fortin, J.-P.; Cullen, N.; Sheline, Y.I.; Taylor, W.D.; Aselcioglu, I.; Cook, P.A.; Adams, P.; Cooper, C.; Fava, M.; McGrath, P.J.; et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage* **2018**, *167*, 104–120. [[CrossRef](#)] [[PubMed](#)]
45. Crombé, A.; Kind, M.; Fadli, D.; le Loarer, F.; Italiano, A.; Buy, X.; Saut, O. Intensity harmonization techniques influence radiomics features and radiomics-based predictions in sarcoma patients. *Sci. Rep.* **2020**, *10*, 15496. [[CrossRef](#)]
46. Lucia, F.; Visvikis, D.; Vallières, M.; Desseroit, M.-C.; Miranda, O.; Robin, P.; Bonaffini, P.A.; Alfieri, J.; Masson, I.; Mervoyer, A.; et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 864–877. [[CrossRef](#)] [[PubMed](#)]
47. Shiri, I.; Amini, M.; Nazari, M.; Hajianfar, G.; Haddadi Avval, A.; Abdollahi, H.; Oveisi, M.; Arabi, H.; Rahmim, A.; Zaidi, H. Impact of feature harmonization on radiogenomics analysis: Prediction of EGFR and KRAS mutations from non-small cell lung cancer PET/CT images. *Comput. Biol. Med.* **2022**, *142*, 105230. [[CrossRef](#)] [[PubMed](#)]
48. Masson, I.; Da-Ano, R.; Lucia, F.; Doré, M.; Castelli, J.; Goislard de Monsabert, C.; Ramée, J.-F.; Sellami, S.; Visvikis, D.; Hatt, M.; et al. Statistical harmonization can improve the development of a multicenter CT-based radiomic model predictive of nonresponse to induction chemotherapy in laryngeal cancers. *Med. Phys.* **2021**, *48*, 4099–4109. [[CrossRef](#)] [[PubMed](#)]
49. Gallardo-Estrella, L.; Lynch, D.A.; Prokop, M.; Stinson, D.; Zach, J.; Judy, P.F.; van Ginneken, B.; van Rikxoort, E.M. Normalizing computed tomography data reconstructed with different filter kernels: Effect on emphysema quantification. *Eur. Radiol.* **2016**, *26*, 478–486. [[CrossRef](#)] [[PubMed](#)]
50. Mackin, D.; Fave, X.; Zhang, L.; Fried, D.; Yang, J.; Taylor, B.; Rodriguez-Rivera, E.; Dodge, C.; Jones, A.K.; Court, L. Credence cartridge radiomics phantom CT scans—The cancer imaging archive (TCIA) public access—Cancer imaging archive wiki. *Cancer Imaging Arch.* **2017**. [[CrossRef](#)]
51. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The cancer imaging archive (TCIA): Maintaining and operating a public information repository. *J. Digit. Imaging* **2013**, *26*, 1045–1057. [[CrossRef](#)] [[PubMed](#)]
52. Van Griethuysen, J.J.M.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.H.; Fillion-Robin, J.-C.; Pieper, S.; Aerts, H.J.W.L. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)] [[PubMed](#)]
53. Team, R.C. *R Language Definition*; R Foundation for Statistical Computing: Vienna, Austria, 2000.
54. Gandrud, C. *Reproducible Research with R and R Studio*; CRC Press: Boca Raton, FL, USA, 2013; ISBN 9781466572843.
55. Lin, L.I. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **1989**, *45*, 255–268. [[CrossRef](#)] [[PubMed](#)]



56. Stevenson, M.; Stevenson, M.M.; BiasedUrn, I. Package “epiR”. Available online: <https://vps.fmvz.usp.br/CRAN/web/packages/epiR/epiR.pdf> (accessed on 15 January 2022).
57. McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **1947**, *12*, 153–157. [[CrossRef](#)]
58. Denzler, S.; Vuong, D.; Bogowicz, M.; Pavic, M.; Frauenfelder, T.; Thierstein, S.; Eboulet, E.I.; Maurer, B.; Schniering, J.; Gabrys, H.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. [[CrossRef](#)]
59. 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. [[CrossRef](#)]
60. Ibrahim, A.; Barufaldi, B.; Refaee, T.; Silva Filho, T.M.; Acciavatti, R.J.; Salahuddin, Z.; Hustinx, R.; Mottaghy, F.M.; Maidment, A.D.A.; Lambin, P. MaasPenn radiomics reproducibility score: A novel quantitative measure for evaluating the reproducibility of CT-based handcrafted radiomic features. *Cancers* **2022**, *14*, 1599. [[CrossRef](#)]
61. Li, Y.; Ammari, S.; Balleyguier, C.; Lassau, N.; Chouzenoux, E. Impact of preprocessing and harmonization methods on the removal of scanner effects in brain MRI radiomic features. *Cancers* **2021**, *13*, 3000. [[CrossRef](#)]
62. Lupean, R.-A.; Ștefan, P.-A.; Csutak, C.; Lebovici, A.; Măluțan, A.M.; Buiga, R.; Melincovici, C.S.; Mișu, C.M. Differentiation of endometriomas from ovarian hemorrhagic cysts at magnetic resonance: The role of texture analysis. *Medicina* **2020**, *56*, 487. [[CrossRef](#)]