

## Article

# Reproducibility of CT-Based Hepatocellular Carcinoma Radiomic Features across Different Contrast Imaging Phases: A Proof of Concept on SORAMIC Trial Data

Abdalla Ibrahim <sup>1,2,3,4,\*</sup>, Yousif Widaatalla <sup>1,†</sup>, Turkey Refaee <sup>1,5,\*</sup>, Sergey Primakov <sup>1,4</sup>, Razvan L. Miclea <sup>2</sup>, Osman Öcal <sup>6</sup>, Matthias P. Fabritius <sup>6</sup>, Michael Ingrisch <sup>6</sup>, Jens Ricke <sup>6</sup>, Roland Hustinx <sup>3</sup>, Felix M. Mottaghy <sup>2,4</sup>, Henry C. Woodruff <sup>1,2</sup>, Max Seidensticker <sup>6,‡</sup> and Philippe Lambin <sup>1,2,‡</sup>

- <sup>1</sup> The D-Lab, Department of Precision Medicine, GROW—School for Oncology, Maastricht University, 6200 MD Maastricht, The Netherlands; y.widaatalla@maastrichtuniversity.nl (Y.W.); S.primakov@maastrichtuniversity.nl (S.P.); h.woodruff@maastrichtuniversity.nl (H.C.W.); philippe.lambin@maastrichtuniversity.nl (P.L.)
  - <sup>2</sup> Department of Radiology and Nuclear Medicine, Maastricht University Medical Centre+, 6200 MD Maastricht, The Netherlands; razvan.miclea@mumc.nl (R.L.M.); felix.mottaghy@mumc.nl (F.M.M.)
  - <sup>3</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@ulg.ac.be
  - <sup>4</sup> Department of Nuclear Medicine and Comprehensive Diagnostic Center Aachen (CDCA), University Hospital RWTH Aachen University, 52074 Aachen, Germany
  - <sup>5</sup> Department of Diagnostic Radiology, Faculty of Applied Medical Sciences, Jazan University, Jazan 45142, Saudi Arabia
  - <sup>6</sup> Department of Radiology, University Hospital, LMU Munich, 80336 Munich, Germany; osman.oecal@med.uni-muenchen.de (O.Ö.); Matthias.Fabritius@med.uni-muenchen.de (M.P.F.); Michael.Ingrisch@med.uni-muenchen.de (M.I.); jens.ricke@med.uni-muenchen.de (J.R.); Max.Seidensticker@med.uni-muenchen.de (M.S.)
- \* Correspondence: a.ibrahim@maastrichtuniversity.nl (A.I.); t.refaee@maastrichtuniversity.nl (T.R.)  
† These authors share equal contribution.  
‡ These authors share last authorship.



**Citation:** Ibrahim, A.; Widaatalla, Y.; Refaee, T.; Primakov, S.; Miclea, R.L.; Öcal, O.; Fabritius, M.P.; Ingrisch, M.; Ricke, 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. <https://doi.org/10.3390/cancers13184638>

Received: 17 August 2021

Accepted: 13 September 2021

Published: 16 September 2021

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



**Copyright:** © 2021 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/>).

**Simple Summary:** Radiomics has been reported to have potential for correlating with clinical outcomes. However, handcrafted radiomic features (HRFs)—the quantitative features extracted from medical images—are limited by their sensitivity to variations in scanning parameters. Furthermore, radiomics analyses require big data with good quality to achieve desirable performances. In this study, we investigated the reproducibility of HRFs between scans acquired with the same scanning parameters except for the imaging phase (arterial and portal venous phases) to assess the possibilities of merging scans from different phases or replacing missing scans from a phase with other phases to increase data entries. Additionally, we assessed the potential of ComBat harmonization to remove batch effects attributed to this variation. Our results show that the majority of HRFs were not reproducible between the arterial and portal venous phases before or after ComBat harmonization. We provide a guide for analyzing scans of different imaging phases.

**Abstract:** Handcrafted radiomic features (HRFs) are quantitative imaging features extracted from regions of interest on medical images which can be correlated with clinical outcomes and biologic characteristics. While HRFs have been used to train predictive and prognostic models, their reproducibility has been reported to be affected by variations in scan acquisition and reconstruction parameters, even within the same imaging vendor. In this work, we evaluated the reproducibility of HRFs across the arterial and portal venous phases of contrast-enhanced computed tomography images depicting hepatocellular carcinomas, as well as the potential of ComBat harmonization to correct for this difference. ComBat harmonization is a method based on Bayesian estimates that was developed for gene expression arrays, and has been investigated as a potential method for harmonizing HRFs. Our results show that the majority of HRFs are not reproducible between the arterial and portal venous imaging phases, yet a number of HRFs could be used interchangeably

between those phases. Furthermore, ComBat harmonization increased the number of reproducible HRFs across both phases by 1%. Our results guide the pooling of arterial and venous phases from different patients in an effort to increase cohort size, as well as joint analysis of the phases.

**Keywords:** hepatocellular carcinoma; CT radiomics; domain translation; reproducibility

## 1. Introduction

Recent decades have witnessed vast advances in computational power, artificial intelligence, and medical imaging techniques [1], which have provided a unique opportunity for transforming the abundant amounts of medical imaging into mineable quantitative data. This concept acquired much scientific attention recently, and a branch of medical imaging analysis—known as handcrafted radiomics—emerged as a result [2]. Handcrafted radiomic features (HRFs) are quantitative features extracted with high throughput from medical imaging, with its varying modalities. The hypothesis is that medical images carry more data than can be seen by trained human eyes, and that these data can be decoded using HRFs; i.e., correlations between HRFs and underlying biology could potentially exist [3]. Since the introduction of the field, many studies have reported on the potential of radiomic signatures to predict clinical endpoints, the majority of which were performed on computed tomography (CT) [4–7], magnetic resonance (MR) [8–10], and positron emission tomography (PET) scans [11,12].

Hepatocellular carcinoma (HCC) is the most common primary liver cancer, the fifth most common malignancy worldwide, and a leading cause of cancer-related mortality [13]. Different diagnostic approaches and treatment modalities are used clinically depending on the characteristics of the patient and the progression of the disease [14,15]. Contrast-enhanced computed tomography (CE-CT) scans are considered one of the main diagnostic tools for HCC. CE-CT can be acquired at different times following the injection of the contrast agent to acquire arterial, venous or late phase scans. Each phase shows specific characteristics for HCC lesions. However, there is still a clinical need for reliable non-invasive tools that could aid in diagnosing and devising individualized treatment plans for HCC patients. Several studies have investigated and reported on the potential of HRFs to aid clinical decision making in HCC patients [16–19].

While numerous studies have reported on the potential of HRFs in aiding clinical decision making on HCC and other diseases, several hurdles hindering the clinical translation of radiomic signatures to clinical decision support systems have been identified. These hurdles include the reproducibility of HRFs in test-retest studies, their sensitivity to variations in acquisition and reconstruction parameters of the scans, inter-observer variability, and the need for big data [20–26]. However, the need for big data in radiomics analysis necessitates the exploration of methods for combining and comparing retrospective medical imaging databases.

A number of studies have tried to address the issue of reproducibility of HRFs using ComBat harmonization [27–30]. ComBat harmonization is a method that was developed to remove the batch effects in gene expression arrays [31]. The studies that investigated the application of ComBat in radiomics analyses reported on the improvement in performance metrics of developed radiomic signatures after the application of ComBat compared to before, and recommended the use of the method. Other studies investigated the reproducibility of HRFs on phantom datasets acquired with different settings [32], or with a single parameter difference [33], and reported that the performance of ComBat is dependent on the data under study. These studies recommended a framework to assess the reproducibility of HRFs. However, to date, no study has reported on the agreement in HRFs across different phases, the potential of ComBat to remove the effects of different imaging phases from HRFs, which could allow the proper combination of phases in a single analysis, or the interchangeability of HRFs across phases to allow the use of different

imaging scans per patient. Furthermore, no study has performed a reproducibility analysis for HRFs following ComBat harmonization on patients' scans acquired with a single parameter difference.

We hypothesize that the time of acquisition after the injection of the contrast agents adds another level of complexity to be accounted for in the radiomics analysis, as HRFs might be affected by the appearance of contrast due to the variations in the distribution of the contrast within the lesions. As a proof of concept, we investigate the sensitivity of HRFs extracted from CE-CT scans depicting HCC acquired during the arterial and portal venous phases when all other acquisition and reconstruction parameters were fixed. Furthermore, we investigate the potential of the ComBat harmonization for domain translation of the HRFs extracted from these scans. Ultimately, we aim to (i) guide the identification of HRFs that can be used interchangeably between arterial and venous phase scans, which could increase the number of scans that can be included in a CE-CT based radiomics study; and (ii) identify the features that can be used in studies analyzing both phases simultaneously to maximize the information extracted from ROIs.

## 2. Materials and Methods

### 2.1. Patients and Imaging Data

The imaging data were originally collected for the European multicenter clinical trial (SORAMIC) [34]. Imaging data for 424 patients diagnosed with HCC (using cyto-histological criteria, radiologic criteria, or a combination of both) were obtained for the SORAMIC trial, of which 338 scans were available for analysis in this study. Scans that contained artifacts were considered of poor quality ( $n = 48$ ). From the available 338 patients with both arterial and portal venous scans available, patients with scans that had any difference in the acquisition or reconstruction parameters, or lacked segmentations reviewed by an expert, were excluded. A total of 61 patients with 104 distinct lesions were finally included in this study (Figure 1). Scans included were acquired from different hospitals using different vendors and protocols. In total, 9 scanner models from 4 different imaging vendors, and a range of scanning parameters, were included, as shown in Table 1. The imaging analysis was approved by the University of Magdeburg institutional review board (IRB00006099, EudraCT no 2009-012576-27), and informed consent was obtained from all included patients. All methods were carried out in accordance with the relevant guidelines and regulations [35].

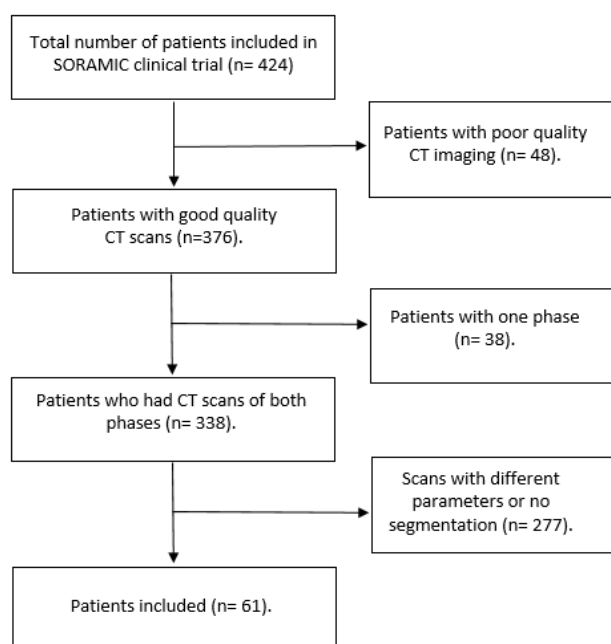


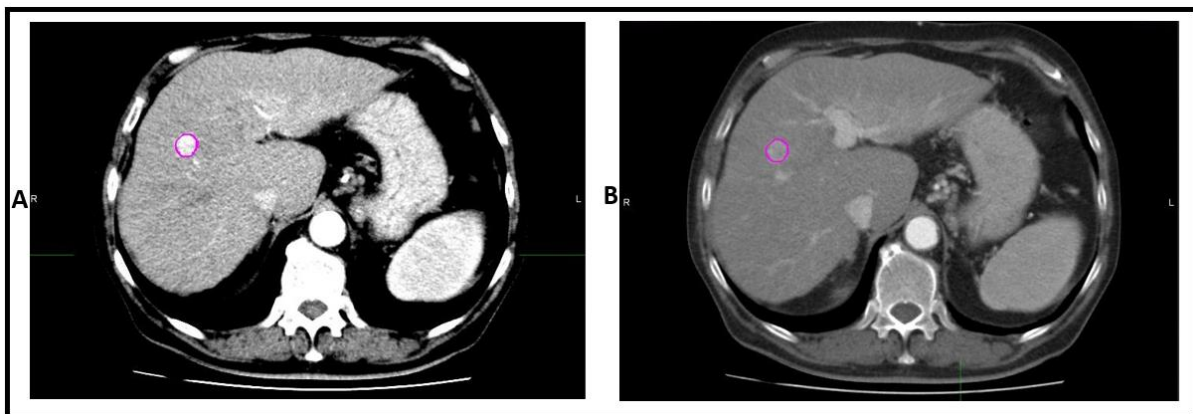
Figure 1. A flowchart showing the patient selection process.

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

Manufacturer	Scanner Model	X-ray Tube Current (kV)	Exposure (mAs)	Convolution Kernels	Slice Thickness (mm)	Pixel Spacing (mm <sup>2</sup> )
TOSHIBA	Aquilion	50–360	2–300	FC13	1–5	0.39 × 0.39 – 0.98 × 0.98
	Aquilion PRIME					
Philips	Brilliance 64			B		
GE	Discovery CT750 HD			STANDARD		
	Optima CT660					
SIEMENS	Sensation 16			B31f		
	SOMATOM Definition AS					
	SOMATOM Definition Flash			I30f, I40f		
	SOMATOM Force			Br40d		

### 2.2. Segmentation and HRFs Extraction

The scans of a single patient were co-registered. The region of interest (ROI) was segmented on each scan while viewing both phases simultaneously and saved to both scans (Figure 2). The segmentations were performed using MIM software (MIM Software Inc., Cleveland, OH) by a medical doctor (Y.W.) with 2 years of experience in image segmentation, and revised by a radiologist (R.M.) with 15 years of experience in medical radiology.



**Figure 2.** An example of ROI segmented in (A) the arterial phase and (B) the portal venous phase.

HRFs were extracted from these ROIs using the software RadiomiX Discovery Toolbox (version, October 2019; <https://www.radiomics.bio>, accessed on 12 January 2021), which calculates HRFs compliant with the Imaging Biomarkers Standardization Initiative (IBSI) [36], in addition to others. Image intensities were binned with a binwidth of 25 Hounsfield Units (HUs) in order to reduce noise levels and to reduce texture matrix sizes, and therewith computation power, with no resampling or further preprocessing of the images. The description of the extracted HRFs was published previously [24].

### 2.3. ComBat Harmonization

The ComBat method employs empirical Bayes to estimate the effects of assigned batches on the data being harmonized. For HRFs, ComBat assumes that a feature value can be approximated by the equation:

$$Y_{ij} = \alpha + \beta X_{ij} + \gamma_i + \delta_i \varepsilon_{ij} \quad (1)$$

where  $\alpha$  is the average value for HRF  $Y_{ij}$  for ROI  $j$  on scanner  $i$ ;  $X$  is a design matrix of the biologic covariates that are known to affect the value of HRFs;  $\beta$  is the vector of regression coefficients corresponding to each biologic covariate;  $\gamma_i$  is the additive effect of scanner  $i$  on HRFs,  $\delta_i$  is the multiplicative scanner effect, and  $\varepsilon_{ij}$  is an error term, presupposed to be

normally distributed with zero mean. Based on the values estimated, ComBat performs feature transformation as given by the formula:

$$Y_{ij}^{ComBat} = \frac{(Y_{ij} - \hat{\alpha} - \hat{\beta}X_{ij} - \gamma_i^*)}{\delta_i^*} + \hat{\alpha} + \hat{\beta}X_{ij} \quad (2)$$

where  $\hat{\alpha}$  and  $\hat{\beta}$  are estimators of the parameters  $\alpha$  and  $\beta$ , respectively; and  $\gamma_i^*$  and  $\delta_i^*$  are the empirical Bayes estimates for the parameters  $\gamma_i$  and  $\delta_i$ , respectively.

#### 2.4. Statistical Analysis

All statistical analyses were performed using R language [37] on RStudio (V 3.6.3) [38]. To determine the reproducibility of HRFs, the concordance correlation coefficient (CCC) between the HRF values across the two phases was calculated [39] using the epiR package [40]. The CCC measures how concordant the values of a given HRF are, as well as the rank of each data point relative to the rest in each batch. HRFs with CCC > 0.9 were considered reproducible and could be interchangeably used between the arterial and venous phase CT scans.

To assess the performance of ComBat, shape features and HRFs with (near) zero variance (HRFs that have the same value in 95% or more of the observations) were removed. The phase of the scan was assigned as the batch for ComBat harmonization. The CCC was calculated after ComBat application and the cutoff of CCC > 0.9 was applied to select the concordant HRFs. The correlation of concordant features with volume was assessed using Pearson correlation. Features that had a correlation coefficient > 0.85 were considered highly correlated. The analysis code used in this study can be found on: (<https://github.com/Abdallahbrahim/The-reproducibility-and-ComBatability-of-Radiomic-features>, accessed on 6 September 2021).

### 3. Results

#### 3.1. Patient Characteristics

The patients included (n = 61) had a median age of 66 years, were mainly male (n = 50, 81.9%) with cirrhotic livers (n = 56, 91.8%), and a minority (n = 11, 18.1%) had portal vein invasion. For more patient characteristics, see Table 2.

#### 3.2. Extracted HRFs

A total of 167 original HRFs were extracted from each of the available 104 ROIs. These HRFs are divided into 11 feature families: Fractal (n = 3), Gray Level Co-occurrence Matrix (GLCM; n = 26), Gray Level Distance Zone Matrix (GLDZM; n = 16), Gray Level Run Length Matrix (GLRLM; n = 15), Gray Level Size Zone Matrix (GLSZM, n = 16), Intensity Histogram (IH; n = 25), Local Intensity (LocInt, n = 2), Neighbouring Gray Level Dependence Matrix (NGLDM; n = 17), Neighbouring Gray Tone Difference Matrix (NGTDM, n = 5), Shape (n = 23), and Statistics (Stats, n = 19).

#### 3.3. The Effects of Differences in Imaging Phase on the Reproducibility of HRFs

Out of the 167 extracted HRFs, 42 (25%) were reproducible (had a CCC > 0.9) across both phases (Figure 3a, shape features were not included to ease the comparison between figures). These HRFs were divided into Shape (n = 22), NGTDM (n = 1), NGLDM (n = 4), IH (n = 2), GLSZM (n = 4), GLRLM (n = 2) and GLDZM (n = 7). The remaining HRFs had a CCC ranging from −0.07 and 0.85, with a median of 0.39.

**Table 2.** Patient characteristics.

Characteristic	N = 61
Gender, male (%)	50 (81.9%)
Age, median (range)	66 (48–81)
Cirrhosis, yes (%)	56 (91.8%)
Child–Pugh grade	
A	56 (91.8%)
B	5 (8.2%)
Diameter of largest lesion, in mm, median (range)	37 (10–220)
Portal vein invasion, yes (%)	11 (18.1%)
Extrahepatic disease yes (%)	7 (11.4%)
* BCLC staging	
A	22 (36.1%)
B	22 (36.1%)
C	17 (27.8%)
** ECOG performance	
0	58 (95.1%)
1	3 (4.9%)

\* Barcelona clinic liver cancer (BCLC) staging; \*\* European Cooperative Oncology Group (ECOG) performance.

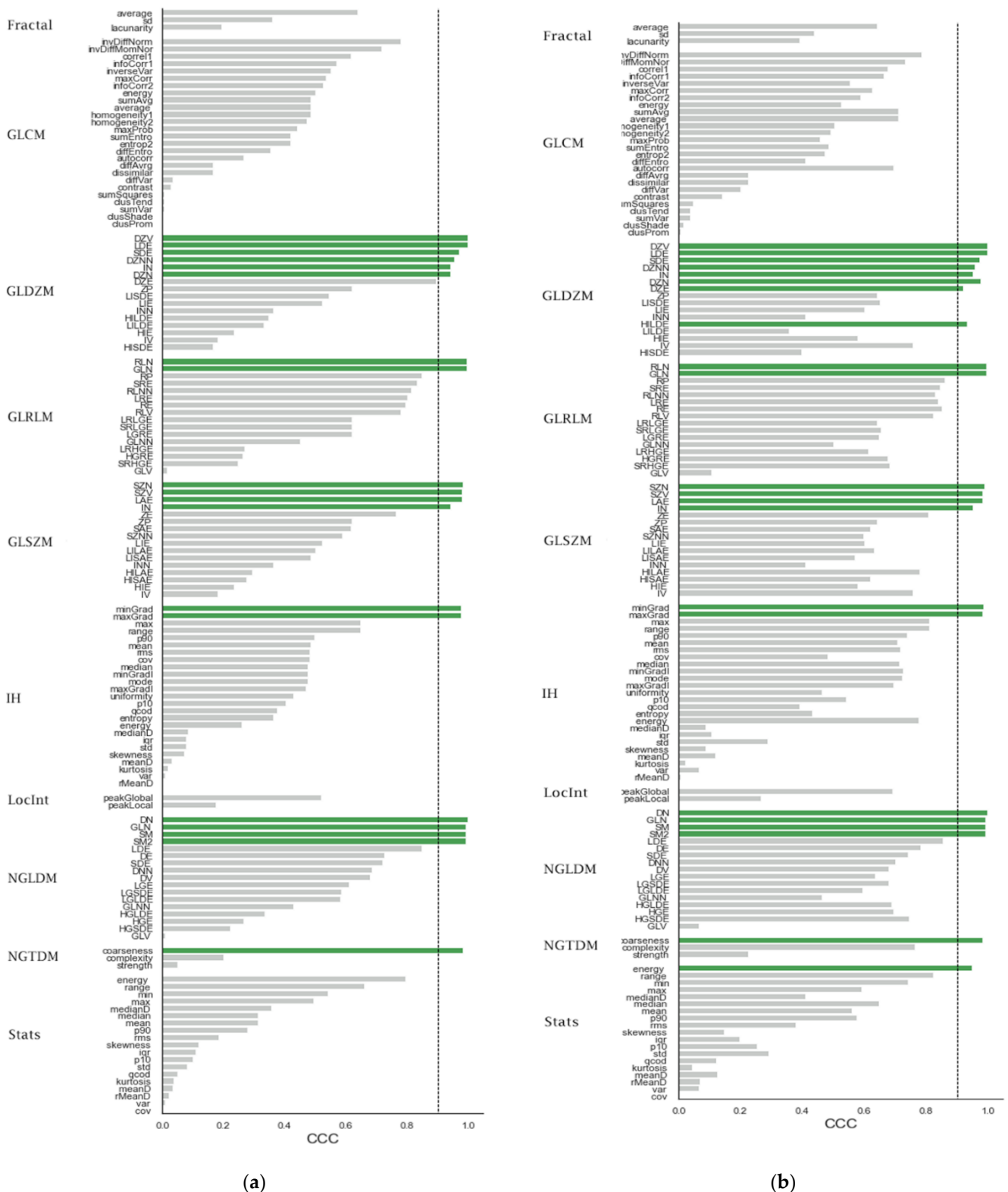
Of the concordant 22 shape features, 8 features were highly correlated with volume ( $R > 0.85$ ), in addition to 1 feature from the NGLDM group (NGLDM\_DN) and 2 features from the GLRLM group (GLRLM\_RLN and GLRLM\_GLN). The remaining features (31, 73.8%) had a correlation coefficient  $< 0.85$ .

### 3.4. The Effects of ComBat on the Reproducibility of HRFs

The application of ComBat harmonization to remove the batch effects attributed to the difference in time between contrast injection and scan acquisition resulted in a total of 44 (26.1%) reproducible HRFs; i.e., 2 extra HRFs became concordant following the application of ComBat: Stats\_energy and GLDZM\_HILDE (Figure 3b). The remaining 20 HRFs had a CCC  $> 0.9$  before and after ComBat harmonization, in addition to the shape features ( $n = 22$ ). The CCC of stats\_energy increased from 0.8 to 0.95 following ComBat harmonization, and that of GLDZM\_HILDE increased from 0.34 to 0.93.

The impact of ComBat on the CCC values had a wide range: 6 HRFs had an increment in CCC between 0.5 and 0.6; 42 HRFs had an increment in CCC between 0.1 and 0.49; 87 HRFs had an increment between 0 and 0.09; and 33 HRFs had a decrement in CCC between  $-0.001$  and  $-0.06$ . Following ComBat harmonization, the number of highly correlated features with volume increased by one feature (Stats\_energy). The concordant features before domain translation maintained their correlation with volume.





**Figure 3.** (a) The CCC values for the different HRFs before ComBat harmonization. (b) The CCC values for the different HRFs after ComBat harmonization.

#### 4. Discussion

In this study, we investigated the reproducibility of HCC CT-based HRFs across the arterial and portal venous imaging phases when all other scanning parameters were

fixed, and whether ComBat harmonization improves the reproducibility of HRFs in such a scenario. Uniquely, this is the first manuscript to investigate the potential of ComBat to remove batch effects attributed to the differences in imaging phase from patient data with a single parameter difference between the compared/harmonized scans. Our results show that the majority of HRFs were significantly affected by the difference in imaging phases, and only a quarter of the total extracted number of HRFs were reproducible across both phases. Moreover, ComBat harmonization did not successfully harmonize the majority of HRFs, even though the differences between the batches compared were limited to the variations in imaging phase.

HRFs are calculated using mathematical formulas applied on the array of values representing the medical image [41]. Changes in the value of units in this array are expected to have an impact on the value calculated by the same formula. Therefore, changes in the scanning parameters are expected to affect the reproducibility of different HRFs variably. Aside from HRFs that are not reproducible in test-retest studies, the sensitivity of the remaining HRFs to the imaging phase can be justified by the increased radio-opacity and the resulting perfusion patterns of contrast within the ROI, and thus, changes in the image array values based on which the HRFs are calculated. As expected, statistics and intensity histogram features, which are simple HRFs based on a single voxel value (e.g., minimum or maximum intensity value) or the description of their distribution (e.g., mean or median intensity value), were found to be the most significantly affected families. On the other hand, also according to expectations, HRFs that do not depend on the intensity values, but the shape of the segmentation (shape features), were found to be reproducible across both phases, with the exception of the shape feature centroid distance, which is based on the distribution of intensity values around the geometric center of the ROI. The copying of segmentations and the inclusion of scans that were acquired identically in both phases allowed isolation of the effects of differences in imaging phases on HRFs. However, in scenarios where acquisition and/or reconstruction parameters, or the segmentation of the ROI changes, the reproducibility of HRFs is expected to be further impacted. This is also in line with what has been reported in a study that investigated the reproducibility of liver parenchyma and tumor HRFs extracted from two contrast-enhanced scans (one phase) taken within a 14 day interval [42]. Therefore, the reproducibility analysis based on the data under study should be an integral part of each radiomics study.

Our study sheds the light on the methodology of combining HRFs from different modalities, either for the purpose of combining different phases/modalities per patient or for combining different phases for different patients. For merging different modalities per patient, we show that a number of HRFs are reproducible across phases. Therefore, models that try to combine different imaging phases per patient are recommended to define which reproducible (test-retest) HRFs vary across the available phases and preselect those for further analysis. Another implication of our findings is allowing the combination of different imaging phases per patient (e.g., due to the lack of data) when only the reproducible HRFs across phases are extracted and compared between the different patients, regardless of the available imaging phase for each patient. This approach can significantly increase the number of data points in retrospective radiomics studies.

The correlation of radiomic features with the volume of the ROI has been considered one of the major points to be assessed in radiomics analysis, since some of the features were reported previously to be surrogates of volume [43]. In our analysis, we observed that the majority of the features identified as concordant (or domain-translatable with ComBat) between the arterial and venous CT scans was considerable, most of which were shape features. However, the majority of features were not found to be highly correlated with volume, which means that these features can decode additional information about the ROIs being investigated.

The number of features that had a CCC value higher than 0.9 was slightly higher after the application of ComBat on the HRFs extracted from the arterial and portal venous phases. ComBat successfully harmonized two additional HRFs compared to the number of



concordant HRFs before domain translation. The majority of HRFs were not concordant across the phases even after the application of ComBat harmonization. The differences in ComBat performance per HRF (and feature families) are also expected, as in contrast to gene expression arrays, HRFs have different levels of complexity and are not expected to be uniformly affected by the batch defined for domain translation. The variant performance of ComBat on HRFs could be explained by the differences in the complexity of HRFs compared to gene expression arrays [21]. The findings are in line with the reproducibility studies that assessed the performance of ComBat on phantom scans, which reported that ComBat harmonization does not successfully harmonize all HRFs, and that its performance is dependent on variations between batches [32,33]. As a consequence, we recommend that the application of ComBat harmonization on HRFs follows a reproducibility analysis with reference values to assess its performance, as it is expected to vary with variations in the dataset batches being harmonized [21]. Other deep learning-based harmonization methods that have been recently investigated [44–47] might be more suitable for domain translation of images acquired in different phases. However, this is yet to be investigated.

While this study provides a proof of concept for the combination/replacement of different imaging phases, we speculate that the set of reproducible HRFs identified in this study is limited to HCC lesions extracted from scans acquired similarly to our dataset. Furthermore, the changes in reconstruction parameters (and sometimes acquisition parameters) between the two imaging phases in clinical routine significantly lowered the number of available scans to perform this analysis. Lastly, the reproducibility of the identified HRFs has to be investigated across different acquisition and reconstruction parameters. However, due to the lack of data, this was not performed. Nevertheless, this study serves as a guide for selecting and/or harmonizing reproducible HRFs in future radiomic studies that utilize contrast-enhanced imaging.

## 5. Conclusions

The majority of HRFs are significantly affected by changes in the imaging phase of the scan. Studies that investigate the potential of combining HRFs from different imaging phases or modalities must investigate the reproducibility and interoperability of the HRFs across the investigated phases for the lesions of interest. Furthermore, a number of HRFs can be interchangeably used between the arterial and portal venous phases, and these can be used to increase data points in retrospective imaging studies. ComBat harmonization increased the number of comparable CT-based HRFs across the arterial and portal venous imaging phases for HCC lesions by 1% in our dataset.

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

**Funding:** The authors acknowledge financial support from the ERC advanced grant (ERC-ADG-2015 n° 694812-Hypoximmuno). The authors also acknowledge financial support from the European Union’s Horizon 2020 research and innovation programme under grant agreement: ImmunoSABR n° 733008, MSCA-ITN-PREDICT n° 766276, CHAIMELEON n° 952172, EuCanImage n° 952103, TRANSCAN Joint Transnational Call 2016 (JTC2016 CLEARLY n° UM 2017-8295). The authors further acknowledge financial support by the Dutch Cancer Society (KWF Kankerbestrijding), project number 12085/2018-2, and Maastricht-Liege Imaging Valley grant, project no. “DEEP-NUCLE”.

**Institutional Review Board Statement:** The imaging analysis was approved by the University of Magdeburg institutional review board (IRB00006099, EudraCT no 2009-012576-27).

**Informed Consent Statement:** Informed consent was obtained from all included patients.

**Data Availability Statement:** The data is privately owned by the trial coordinators.

**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 issue patent on mtDNA (PCT/EP2014/059089) licensed to ptTheragnostic/DNAmito, three non-patented invention (softwares) licensed to ptTheragnostic/DNAmito, Radiomics SA and Health Innovation Ventures and three non-issues, non-licensed patents on Deep Learning-Radiomics and LSRT (N2024482, N2024889, N2024889). He confirms that none of the above entities or funding was involved in the preparation of this paper. Woodruff has minority shares in the company OncoRadiomics. Mottaghy received an advisor fee and reimbursement of travel costs from OncoRadiomics. He reports institutional grants from GE and Nanomab outside the submitted work. The rest of coauthors 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.; et al. Decision Support Systems in Oncology. *JCO Clin. Cancer Inform.* **2019**, *3*, 1–9. [[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](#)]
- Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016**, *278*, 563–577. [[CrossRef](#)]
- Refaee, T.; Wu, G.; Ibrahim, A.; Halilaj, I.; Leijenaar, R.T.H.H.; Rogers, W.; Gietema, H.A.; Hendriks, L.E.L.; Lambin, P.; Woodruff, H.C. The Emerging Role of Radiomics in COPD and Lung Cancer. *Respiration* **2020**, *99*, 99–107. [[CrossRef](#)] [[PubMed](#)]
- Aerts, H.J.W.L. The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review. *JAMA Oncol.* **2016**, *2*, 1636–1642. [[CrossRef](#)] [[PubMed](#)]
- van Timmeren, J.E.; Leijenaar, R.T.H.; van Elmpt, W.; Reymen, B.; Oberije, C.; Monshouwer, R.; Bussink, J.; Brink, C.; Hansen, O.; Lambin, P. Survival prediction of non-small cell lung cancer patients using radiomics analyses of cone-beam CT images. *Radiother. Oncol.* **2017**, *123*, 363–369. [[CrossRef](#)] [[PubMed](#)]
- Panth, K.M.; Leijenaar, R.T.H.; Carvalho, S.; Liewwes, N.G.; Yaromina, A.; Dubois, L.; Lambin, P. Is there a causal relationship between genetic changes and radiomics-based image features? An in vivo preclinical experiment with doxycycline inducible GADD34 tumor cells. *Radiother. Oncol.* **2015**, *116*, 462–466. [[CrossRef](#)] [[PubMed](#)]
- Jethanandani, A.; Lin, T.A.; Volpe, S.; Elhalawani, H.; Mohamed, A.S.R.; Yang, P.; Fuller, C.D. Exploring Applications of Radiomics in Magnetic Resonance Imaging of Head and Neck Cancer: A Systematic Review. *Front. Oncol.* **2018**, *8*, 131. [[CrossRef](#)] [[PubMed](#)]
- Ursprung, S.; Beer, L.; Bruining, A.; Woitek, R.; Stewart, G.D.; Gallagher, F.A.; Sala, E. Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma—a systematic review and meta-analysis. *Eur. Radiol.* **2020**, *30*, 3558–3566. [[CrossRef](#)]
- Samiei, S.; Granzier, R.W.Y.; Ibrahim, A.; Primakov, S.; Lobbes, M.B.I.; Beets-Tan, R.G.H.; van Nijnatten, T.J.A.; Engelen, S.M.E.; Woodruff, H.C.; Smidt, M.L. Dedicated Axillary MRI-Based Radiomics Analysis for the Prediction of Axillary Lymph Node Metastasis in Breast Cancer. *Cancers* **2021**, *13*, 757. [[CrossRef](#)]
- Ibrahim, A.; Vallières, M.; Woodruff, H.; Primakov, S.; Beheshti, M.; Keek, S.; Refaee, T.; Sanduleanu, S.; Walsh, S.; Morin, O.; et al. Radiomics Analysis for Clinical Decision Support in Nuclear Medicine. *Semin. Nucl. Med.* **2019**, *49*, 438–449. [[CrossRef](#)]
- Lovinfosse, P.; Visvikis, D.; Hustinx, R.; Hatt, M. FDG PET radiomics: A review of the methodological aspects. *Clin. Transl. Imaging* **2018**, *6*, 379–391. [[CrossRef](#)]
- Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)]
- Aubé, C.; Oberti, F.; Lonjon, J.; Pageaux, G.; Seror, O.; N’Kontchou, G.; Rode, A.; Radenne, S.; Cassinotto, C.; Vergniol, J.; et al. EASL and AASLD recommendations for the diagnosis of HCC to the test of daily practice. *Liver Int.* **2017**, *37*, 1515–1525. [[CrossRef](#)]
- Finn, R.S.; Qin, S.; Ikeda, M.; Galle, P.R.; Ducreux, M.; Kim, T.-Y.; Kudo, M.; Breder, V.; Merle, P.; Kaseb, A.O. Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *N. Engl. J. Med.* **2020**, *382*, 1894–1905. [[CrossRef](#)] [[PubMed](#)]
- Mokrane, F.-Z.; Lu, L.; Vasseur, A.; Ota, P.; Peron, J.-M.; Luk, L.; Yang, H.; Ammari, S.; Saenger, Y.; Rousseau, H.; et al. Radiomics machine-learning signature for diagnosis of hepatocellular carcinoma in cirrhotic patients with indeterminate liver nodules. *Eur. Radiol.* **2020**, *30*, 558–570. [[CrossRef](#)]
- Wu, J.; Liu, A.; Cui, J.; Chen, A.; Song, Q.; Xie, L. Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on precontrast magnetic resonance images. *BMC Med. Imaging* **2019**, *19*, 1–11. [[CrossRef](#)]

18. Zhou, Y.; He, L.; Huang, Y.; Chen, S.; Wu, P.; Ye, W.; Liu, Z.; Liang, C. CT-based radiomics signature: A potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma. *Abdom. Radiol.* **2017**, *42*, 1695–1704. [[CrossRef](#)] [[PubMed](#)]
19. Wakabayashi, T.; Ouhmich, F.; Gonzalez-Cabrera, C.; Felli, E.; Saviano, A.; Agnus, V.; Savadjiev, P.; Baumert, T.F.; Pessaux, P.; Marescaux, J.; et al. Radiomics in hepatocellular carcinoma: A quantitative review. *Hepatol. Int.* **2019**, *13*, 546–559. [[CrossRef](#)] [[PubMed](#)]
20. Yip, S.S.F.; Aerts, H.J.W.L. Applications and limitations of radiomics. *Phys. Med. Biol.* **2016**, *61*, R150–R166. [[CrossRef](#)] [[PubMed](#)]
21. Ibrahim, A.; Primakov, S.; Beuque, M.; Woodruff, H.C.; Haililaj, 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* **2021**, *188*, 20–29. [[CrossRef](#)]
22. LaRue, R.T.H.M.; Van Timmeren, J.E.; De Jong, E.E.C.; Feliciani, G.; Leijenaar, R.T.H.H.; Schreurs, W.M.J.; Sosef, M.N.; Raat, F.H.P.J.; Van Der Zande, F.H.R.; Das, M.; et al. Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: A comprehensive phantom study. *Acta Oncol.* **2017**, *56*, 1544–1553. [[CrossRef](#)]
23. Van Timmeren, J.E.; Leijenaar, R.T.H.H.; van Elmpt, W.; Wang, J.; Zhang, Z.; Dekker, A.; Lambin, P. Test–Retest Data for Radiomics Feature Stability Analysis: Generalizable or Study-Specific? *Tomography* **2016**, *2*, 361–365. [[CrossRef](#)]
24. 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](#)]
25. Granzier, R.W.Y.; Verbakel, N.M.H.; Ibrahim, A.; Van Timmeren, J.E.; Van Nijnatten, T.J.A.; Leijenaar, R.T.H.H.; Lobbes, M.B.I.; Smidt, M.L.; Woodruff, H.C. MRI-based radiomics in breast cancer: Feature robustness with respect to inter-observer segmentation variability. *Sci. Rep.* **2020**, *10*, 14163. [[CrossRef](#)]
26. Leijenaar, R.T.H.H.; Carvalho, S.; Velazquez, E.R.; Van Elmpt, W.J.C.; Parmar, C.; Hoekstra, O.S.; Hoekstra, C.J.; Boellaard, R.; Dekker, A.L.A.J.; Gillies, R.J.; et al. Stability of FDG-PET Radiomics features: An integrated analysis of test-retest and inter-observer variability. *Acta Oncol.* **2013**, *52*, 1391–1397. [[CrossRef](#)]
27. 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](#)]
28. 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](#)]
29. Da-Ano, R.; Masson, I.; Lucia, F.; Doré, M.; Robin, P.; Alfieri, J.; Rousseau, C.; Mervoyer, A.; Reinhold, C.; Castelli, J.; et al. Performance comparison of modified ComBat for harmonization of radiomic features for multicenter studies. *Sci. Rep.* **2020**, *10*, 10248. [[CrossRef](#)] [[PubMed](#)]
30. 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 Orhac, 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](#)]
31. Johnson, W.E.; Li, C.; Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **2007**, *8*, 118–127. [[CrossRef](#)]
32. Ibrahim, A.; Refaee, T.; Leijenaar, R.T.H.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](#)] [[PubMed](#)]
33. 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](#)]
34. Ricke, J.; Klumpen, H.J.; Amthauer, H.; Bargellini, I.; Bartenstein, P.; de Toni, E.N.; Gasbarrini, A.; Pech, M.; Peck-Radosavljevic, M.; Popović, P.; et al. Impact of Combined Selective Internal Radiation Therapy and Sorafenib on Survival in Advanced Hepatocellular Carcinoma. *J. Hepatol.* **2019**, *71*, 1164–1174. [[CrossRef](#)] [[PubMed](#)]
35. World Medical Association. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research involving Human Subjects. *JAMA* **2013**, *310*, 2191–2194. [[CrossRef](#)]
36. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.W.L.; Andrearczyk, V.; Apte, A.; Ashrafina, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* **2020**, *295*, 328–338. [[CrossRef](#)]
37. R Core Team. *R Language Definition*; R Foundation for Statistical Computing: Vienna, Austria, 2000.
38. Gandrud, C. *Reproducible Research with R and R Studio*; CRC Press: Boca Raton, FL, USA, 2013; ISBN 9781466572843.
39. Lawrence, I.; Lin, K. A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics* **1989**, *45*, 255–268. [[CrossRef](#)]
40. Stevenson, M.; Stevenson, M.M.; BiasedUrn, I. Package “epiR”. Available online: <https://vps.fmvz.usp.br/CRAN/web/packages/epiR/epiR.pdf> (accessed on 10 March 2021).
41. Aerts, H.J.W.L.; Velazquez, E.R.; Leijenaar, R.T.H.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](#)]

42. Perrin, T.; Midya, A.; Yamashita, R.; Chakraborty, J.; Saidon, T.; Jarnagin, W.R.; Gonen, M.; Simpson, A.L.; Do, R.K.G. Short-term reproducibility of radiomic features in liver parenchyma and liver malignancies on contrast-enhanced CT imaging. *Abdom. Radiol.* **2018**, *43*, 3271–3278. [[CrossRef](#)]
43. Welch, M.L.; McIntosh, C.; Haibe-Kains, B.; Milosevic, M.F.; Wee, L.; Dekker, A.; Huang, S.H.; Purdie, T.G.; O’Sullivan, B.; Aerts, H.J.W.L.; et al. Vulnerabilities of radiomic signature development: The need for safeguards. *Radiother. Oncol.* **2019**, *130*, 2–9. [[CrossRef](#)]
44. Andrearczyk, V.; Depeursinge, A.; Müller, H. Neural network training for cross-protocol radiomic feature standardization in computed tomography. *J. Med. Imaging* **2019**, *6*, 024008. [[CrossRef](#)] [[PubMed](#)]
45. Bashyam, V.M.; Doshi, J.; Erus, G.; Srinivasan, D.; Abdulkadir, A.; Habes, M.; Fan, Y.; Masters, C.L.; Maruff, P.; Zhuo, C.; et al. Medical Image Harmonization Using Deep Learning Based Canonical Mapping: Toward Robust and Generalizable Learning in Imaging. *arXiv* **2020**, arXiv:2010.05355.
46. Modanwal, G.; Vellal, A.; Mazurowski, M.A. Normalization of breast MRIs using cycle-consistent generative adversarial networks. *arXiv* **2019**, arXiv:1912.08061.
47. Dewey, B.E.; Zhao, C.; Reinhold, J.C.; Carass, A.; Fitzgerald, K.C.; Sotirchos, E.S.; Saidha, S.; Oh, J.; Pham, D.L.; Calabresi, P.A.; et al. DeepHarmony: A deep learning approach to contrast harmonization across scanner changes. *Magn. Reson. Imaging* **2019**, *64*, 160–170. [[CrossRef](#)] [[PubMed](#)]