TOMOGRAPHY® RESEARCH ARTICLE

Reliability of Radiomic Features Across Multiple Abdominal CT Image Acquisition Settings: A Pilot Study Using ACR CT Phantom

Lin Lu, Yongguang Liang, Lawrence H. Schwartz, and Binsheng Zhao

Department of Radiology, Columbia University Medical Center, New York, NY

Corresponding Author:
Binsheng Zhao, DSc
Department of Radiology, Columbia University Medical Center,
710 West 168th Street, B26, New York, NY 10032;
E-mail: bz2166@cumc.columbia.edu

Key Words: quantitative imaging biomarkers, Radiomic Features, Reliability, Abdominal CT
Abbreviations: Computed tomography (CT), radiomic features (RFs), gray-level co-occurrence matrix (GLCM), regions of interest (ROIs)

ABSTRACT

We studied the reliability of radiomic features on abdominal computed tomography (CT) images reconstructed with multiple CT image acquisition settings using the ACR (American College of Radiology) CT Phantom. Twenty-four sets of CT images of the ACR CT phantom were attained from a GE Discovery 750HD scanner using 24 different image acquisition settings, combinations of 4 tube currents (25, 50, 100, 200 Effective mAs), 3 slice thicknesses (1.25, 2.5, 5 mm), and 2 convolution kernels (STANDARD and SOFT). Polyethylene (−95 HU) and acrylic (120 HU) of the phantom model were selected for calculating real feature value; a noise-free, computer-generated phantom image series that reproduced the 2 objects and the background was used for calculating reference feature value. Feature reliability was defined as the degree of predicting reference feature value from real feature value. Radiomic features *mean, std, skewness, kurtosis,* gray-level co-occurrence matrix (*GLCM*)*-energy, GLCM-contrast, GLCM-correlation, GLCM-homogeneity* were investigated. The value of $R^2 \geq 0.85$ was considered to be of high reliability. The reliability of *mean* and *std* were high across all image acquisition settings. At 200 Effective mAs, all features except *GLCM-homogeneity* showed high reliability, whereas at 25 Effective mAs, most features (except *mean* and *std*) showed low reliability. From high to low, reliability was ranked in the following order: *mean, std, skewness, kurtosis, GLCM-energy, correlation, contrast* and *homogeneity*. CT image acquisition settings affected the reliability of radiomic features. High reliable features were attained from images reconstructed at high tube current and thick slice thickness.

INTRODUCTION

Medical imaging plays an ever greater role in disease diagnosis and patient care. One of the most exciting new areas related to cancer diagnosis, treatment planning, and response assessment is the field of radiomics, which involves the extraction and analysis of a large number of quantitative imaging features from medical images for characterization of tumor and tissue phenotypes (1, 2).

Owing to the associations between tumor phenotypes and underlying biological processes, radiomic features (RFs) or RF-derived phenotypes can act as biomarkers that convey information about disease to help with the management of therapies. To date, radiomics has shown promise in improving cancer diagnosis and prognostic assessment in several tumor types including lung (3-5), brain (6), breast (7), liver (8-10), kidney (11), and esophagus (12) cancers. Moreover, RFs also exhibit correlations with genetic mutation status (5) and disease recurrence (13), as well as therapeutic response (14) and survival (15) in lung cancer.

While serving as an imaging biomarker for oncology, the influence of image acquisition settings on RFs should be well understood before the biomarker can be fully utilized (16). Until now, numerous studies have been conducted on the "reproducibility" of RFs (17-20), which refers to whether feature values could remain the same when reimaged using different equipment and different image acquisition settings. To the best of our knowledge, with the exception of studies on the accuracy of volume measurements (21, 22), there has been no report to date exploring the "reliability" of RFs. "Reliability" refers to whether true feature value could be maintained when imaged using different scanners and image acquisition settings. The true feature value in our study was defined as the feature value that was calculated on computed tomography (CT) image within which the CT number of each tissue composition was equal to its theoretical CT number at 120 kVp, for example, air equals to −1000 HU, and water equals to 0 HU. Thus, true feature value was also called as reference value in our study.

ISSN 2379-1381 http://dx.doi.org/10.18383/j.tom.2019.00005

Table 1. Image Acquisition Parameters

Scanner	GE Discovery 750HD (64 slices)
kVp	120
Display field of view (cm)	22
Pitch	1.375
Tube Currents (effective mAs)	25, 50, 100, 200
Rotation time (second)	0.7
Beam width (mm)	40 (64x0.625)
Slice thickness (mm)	1.25, 2.5, 5
Overlap (%)	0
Reconstruction algorithms	STANDARD, SOFT

The challenge of such a reliability study lies in the fact that reference values for RFs are generally quite difficult to obtain, especially for in vivo lesions, because of unknown tissue composition, as well as anatomic, physiologic, and even positional variations among different patients. In view of this point, we aimed to carry out a pilot study on RF reliability using the ACR CT phantom (American College of Radiology CT accreditation phantom) (23). The ACR CT phantom is a widely used CT QC phantom, and has a well-defined CT number for each object inside module 1.

In this study, we attained CT images of the phantom under 24 image acquisition settings using a GE Discovery 750HD scanner (GE Healthcare, Waukesha, WI). The reliability of 8 widely used RFs—*mean, std, skewness, kurtosis, GLCM* [*gray-level co-occurrence matrix* (24)]-*energy, GLCM-contrast, GLCM-correlation,* and *GLCM-homogeneity*—was investigated on the 24 sets of CT images.

METHODS

Scanning the ACR CT Phantom

A Gammex CT ACR 464 phantom was scanned on a GE Discovery 750HD scanner using a routine adult abdomen protocol at 4 different tube currents (25, 50, 100, 200 Effective mAs). The CT images were then reconstructed with 3 different slice thicknesses (1.25, 2.5, 5 mm) and 2 convolution kernels (STANDARD, SOFT), resulting in a total of $4 \times 3 \times 2 = 24$ sets of CT images. The CT scanning parameters used in this study are listed in Table 1.

Preparation of Image Region and ROIs for Extracting Real Feature Value

The ACR CT phantom is composed of 4 modules and primarily constructed from water-equivalent materials (23). Each module contains several components made of different materials. In our study, 2 circular objects from module 1, made of polyethylene and acrylic each, were selected to create image patterns for feature extraction. Polyethylene and acrylic are materials with CT numbers of −95 HU and 120 HU at a 120-kVp setting falling within the ranges of the abdominal CT window.

For each object, a 2-dimensional region of 45×45 mm containing the object was cropped from the CT image located at the center of module 1 along the axial direction. Within the cropped region, 100 regions of interest (ROIs) were randomly generated. The criteria to generate ROIs included the following:

(1) The center of the ROI should be located inside the object.
(2) ROI shall cover part of the object and part of the background outside the object, for the purpose of studying radiomic features on nonhomogenous patterns rather than only on homogenous patterns, such as that derived from cartridge phantoms filled with paper/rubber in the literature (19).
(3) The size of the ROI must range from 12×12 mm to 18×18 mm; the sizes of the cropped region and ROIs were empirically determined on the basis of the physical size of the object (a cylinder with diameter = 25 mm and depth = 4 cm as provided in the manual of ACR CT phantom). The process of preparing the cropped region and ROIs is illustrated in Figure 1.

Preparation of Computer-Generated Images for Extracting Reference Feature Value

A noise-free digital image series to simulate module 1 of the ACR CT phantom was generated for the extraction of reference feature values. The 2 selected objects (polyethylene and acrylic) were reproduced via an image-processing algorithm on the basis of designated parameters (eg, location, size, shape, and density in CT number) provided in the phantom manual (23). The rest of the computer-generated images were defined as water-equivalent background with a CT number = 0. The image region and ROIs for feature extraction from the computer-generated images were copied from those used in the scanned phantom images to guarantee that they were identical so that variations introduced by position misalignment and density difference could be minimized and the bias of real value to reference value would be purely because of the different image acquisition settings.

Extraction of Feature

In our study, 8 2D RFs were investigated, including 4 histogram-based features—*mean*, *std (standard deviation)*, *skewness*, and *kurtosis*—and four texture-based GLCM features (24), *GLCM-energy*, *GLCM-contrast*, *GLCM-correlation*, and *GLCM-homogeneity*. *Mean*, *std*, *skewness*, and *kurtosis* are first-order statistic features to characterize an histogram of image intensity. GLCM features are textural features characterizing the gray-tone spatial dependencies of an image, that is, quantifying the relationship between pixels within an ROI. Details of definitions of the 8 RFs are provided in the online Supplemental Material.

In the implementation, the 8 RFs were calculated on each ROI by using an in-house feature extraction algorithm programmed on the MATLAB 2016b platform (MathWorks, Natick, MA). Before feature calculation, images were interpolated into isotropic pixel spacing of 0.5×0.5 mm^2.

Reliability of Feature

In our study, feature reliability was defined as the degree of predicting reference feature value from real feature value. Reference feature value was the feature value extracted from noise-free computer-generated phantom images, while real feature value is the feature value extracted from CT images attained from the physical ACR phantom. High predictability means that

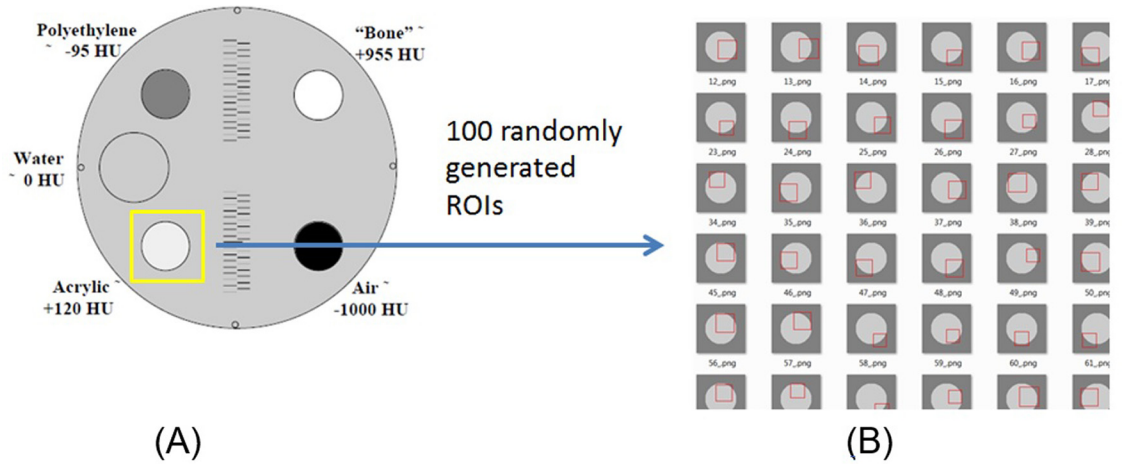

Figure 1. Example of an ACR CT phantom image being prepared for feature extraction. One region containing the object (yellow frame of 45 × 45 mm) is selected from the ACR CT phantom image (A). Samples of 100 randomly generated nonhomogenous ROIs of the object (B). Each ROI was a 2-dimensional square with random location and random size ranging from 12 × 12 mm to 18 × 18 mm.

a change in reference feature value can be correctly reflected by a proportional change in the real feature value. If an RF exhibited high predictability under a certain image acquisition setting, then the RF calculation was believed to be reliable.

Consequently, R^2, a statistical metric widely used to assess the proportion of variance in the dependent variable that is predictable from the independent variable, was adopted to quantify feature reliability. An R^2 value of 1 indicated that the reference feature value could be predicted by a real feature value to a degree of 100%, whereas an R^2 value of 0 indicated that there was no relation between reference feature value and real feature value. The R^2 equation can be defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (x_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (1)$$

Where x_i represents real feature value extracted from the ith ROI on the ACR CT phantom images, y_i represents reference feature value extracted from the corresponding ith ROI on the computer-generated phantom images, $\bar{y}$ represents the mean of reference feature values extracted from 200 ROIs, and n equals 200 (100 ROIs from each of the 2 ACR objects).

Figure 2 shows an example of how to use R^2 to assess feature reliability under certain image acquisition setting. The graphs (A) and (B) in Figure 2 present the skewness values, one of the histogram-based RFs, calculated from ROIs under the image acquisition settings of "convolution kernel = STANDARD, slice thickness = 1.25 mm, and Effective mAs = 200" and "convolution kernel = STANDARD, slice thickness = 1.25 mm, Effective mAs = 25," respectively. The feature data used to estimate R^2 value consisted of 200 pairs of skewness values, corresponding to the reference and real skewness values calculated on the 200 ROIs on the computer-generated and physical phantom images, respectively. As shown in Figure 2, high reliability (R^2 = 0.9575) indicated that reference skewness values approximated the real skewness values measured at high tube current, while low reliability (R^2 = 0.4021) indicated reference skewness values diverged from real skewness values measured at low tube current.

RESULTS

Figure 3 shows the reliability values for the 8 RFs under 24 image acquisition settings, combinations of 4 tube currents (25, 50, 100, 200 Effective mAs), 3 slice thicknesses (1.25, 2.5, 5 mm), and 2 convolution kernels (STANDARD and SOFT). Overall, we were able to observe that feature reliability decreased with a decrease in tube current, features were more reliable on 5-mm CT images than on 1.25- and 2.5-mm CT images, there was little difference in feature reliability between CT images of STANDARD and SOFT convolution kernels, and histogram-based RFs are more reliable than textural RFs.

We averaged the reliability values across individual image acquisition parameters to further investigate their influence (Table 2). For example, when investigating the influence of "200 Effective mAs," we averaged the feature reliability values of "STANDARD_ST125_EffmAs200," "STANDARD_ST250_EffmAs200," "STANDARD_ST500_EffmAs200," "SOFT_ST125_EffmAs200," "SOFT_ST250_EffmAs200," and "SOFT_ST500_EffmAs200" together as presented in Figure 3.

To facilitate the analysis, we empirically set $R^2 > 0.85$ as high reliability. As shown in Table 2, in the case of tube current, 100 Effective mAs could be regarded as a threshold to guide the application of RFs, that is, using tube current ≥100 Effective mAs resulted in more reliable RFs, especially the histogram-based RFs, while using tube current <100 Effective mAs produced only a few reliable RFs. For slice thickness, 5-mm CT images yielded more reliable RFs. For convolution kernel, the STANDARD and SOFT showed similar influence on feature re-

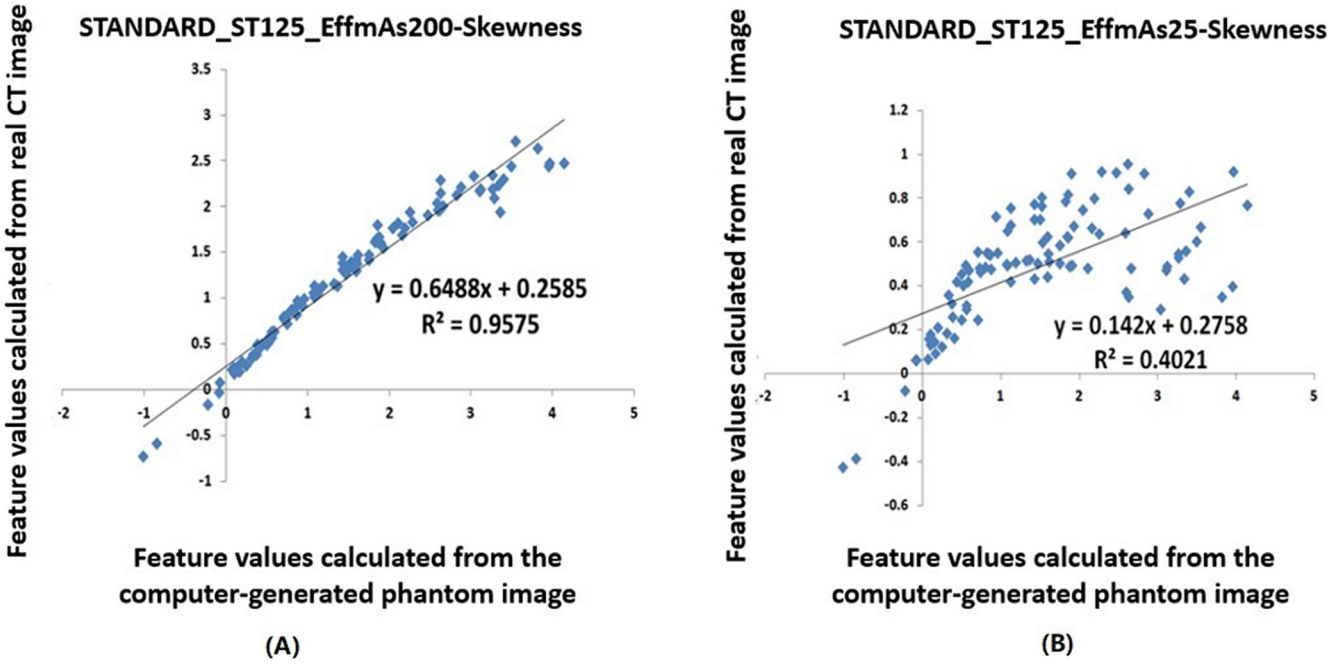

Figure 2. Reliability of the skewness feature at high tube current (A) and low tube current (B), respectively. Each point on the plot corresponds to the values calculated from one randomly generated ROI on the computer-generated (X-axis) and the physical phantom images (Y-axis), respectively. There are a total of 200 points on each plot corresponding to the 200 randomly generated ROIs from the two selected objects in the phantom.

liability. The feature *mean* and *std* showed extremely high reliability across all image acquisition settings.

We observed an obvious unusual trend that the average reliability of GLCM-energy at slice thickness of 1.25 mm was higher than that at slice thickness of 2.5 mm (see Table 2). As we turned to the details of reliability presented in Figure 3, we found that the unusual trend was caused by a great drop of reliability at tube current 25 Effective mAs, a very low dose condition for the slice thickness of 2.5 mm. Actually, based on our results, low tube current easily led to unusual trends for some RFs, for example, GLCM-homogeneity at 50 and 25 Effective mAs and GLCM-homogeneity at 1.25- and 2.5-mm slice thicknesses in Table 2.

Features	STANDARD_ ST125_Effm As200	STANDARD_ ST125_Effm As100	STANDARD_ ST125_Effm As50	STANDARD_ ST125_Effm As25	STANDARD_ ST250_Effm As200	STANDARD_ ST250_Effm As100	STANDARD_ ST250_Effm As50	STANDARD_ ST250_Effm As25	STANDARD_ ST500_Effm As200	STANDARD_ ST500_Effm As100	STANDARD_ ST500_Effm As50	STANDARD_ ST500_Effm As25
mean	0.998	0.999	0.997	0.993	0.996	0.998	0.998	0.996	0.996	0.997	0.999	0.996
std	0.989	0.985	0.965	0.953	0.990	0.987	0.975	0.966	0.990	0.989	0.986	0.985
skewness	0.958	0.812	0.511	0.402	0.973	0.929	0.749	0.454	0.984	0.972	0.907	0.780
kurtosis	0.931	0.824	0.619	0.500	0.969	0.933	0.800	0.565	0.972	0.968	0.927	0.813
glcm-energy	0.941	0.947	0.885	0.768	0.963	0.935	0.930	0.498	0.924	0.947	0.938	0.760
glcm-contrast	0.895	0.656	0.124	0.328	0.917	0.771	0.238	0.119	0.922	0.867	0.619	0.538
glcm-correlation	0.847	0.794	0.802	0.754	0.868	0.811	0.793	0.813	0.894	0.828	0.796	0.814
glcm-homogeneity	0.716	0.546	0.228	0.333	0.746	0.633	0.037	0.168	0.746	0.736	0.296	0.495
Features	**SOFT_ST125 _EffmAs200**	**SOFT_ST125 _EffmAs100**	**SOFT_ST125 _EffmAs50**	**SOFT_ST125 _EffmAs25**	**SOFT_ST250 _EffmAs200**	**SOFT_ST250 _EffmAs100**	**SOFT_ST250 _EffmAs50**	**SOFT_ST250 _EffmAs25**	**SOFT_ST500 _EffmAs200**	**SOFT_ST500 _EffmAs100**	**SOFT_ST500 _EffmAs50**	**SOFT_ST500 _EffmAs25**
mean	0.997	0.999	0.998	0.993	0.996	0.998	0.998	0.995	0.996	0.997	0.999	0.997
std	0.989	0.986	0.972	0.959	0.989	0.988	0.978	0.973	0.989	0.989	0.987	0.985
skewness	0.955	0.883	0.649	0.453	0.980	0.957	0.841	0.602	0.985	0.980	0.944	0.862
kurtosis	0.951	0.882	0.723	0.591	0.972	0.955	0.874	0.657	0.970	0.972	0.953	0.870
glcm-energy	0.926	0.930	0.835	0.786	0.959	0.937	0.830	0.504	0.925	0.952	0.901	0.733
glcm-contrast	0.905	0.748	0.208	0.365	0.923	0.830	0.402	0.205	0.922	0.880	0.745	0.669
glcm-correlation	0.871	0.799	0.813	0.750	0.888	0.825	0.801	0.808	0.910	0.846	0.806	0.814
glcm-homogeneity	0.746	0.507	0.082	0.275	0.817	0.736	0.048	0.165	0.775	0.788	0.309	0.592

Figure 3. Reliability of 8 radiomic features under 24 image acquisition settings. Top panel with pink title: reconstructed using Standard kernel; Bottom panel with yellow title: reconstructed using Soft Kernel. For example, the number of 0.998 in the top-left cell is the R^2 value of the feature *mean* calculated between the computer-generated image and CT scan image obtained at 200 Effective mAs and reconstructed using STANDARD kernel, 1.25 mm slice thickness.

Table 2. Average of Reliability Values Under Individual Image Acquisition Parameters

Features	Tube Current				Slice Thickness			Convolution Kernel		
	200 Effective mAs	100 Effective mAs	50 Effective mAs	25 Effective mAs	1.25 mm	2.5 mm	5.0 mm	STANDARD	SOFT	All
Mean	0.997	0.998	0.998	0.995	0.997	0.997	0.997	0.997	0.997	0.997
Std	0.989	0.987	0.977	0.970	0.975	0.981	0.987	0.980	0.982	0.981
Skewness	0.973	0.922	0.767	0.592	0.703	0.811	0.927	0.786	0.841	0.813
Kurtosis	0.961	0.922	0.816	0.666	0.753	0.841	0.931	0.818	0.864	0.841
GLCM-energy	0.940	0.942	0.887	0.675	0.877	0.819	0.885	0.870	0.852	0.861
GLCM-contrast	0.914	0.792	0.389	0.371	0.529	0.551	0.770	0.583	0.650	0.616
GLCM-correlation	0.880	0.817	0.802	0.792	0.804	0.826	0.839	0.818	0.828	0.823
GLCM-homogeneity	0.758	0.658	0.167	0.338	0.429	0.419	0.592	0.473	0.487	0.480

DISCUSSION

In this study, we introduced the concept of RF reliability and evaluated the RF reliability of 8 commonly used RFs under 24 different image acquisition settings. The 24 image acquisition settings involved 3 image acquisition parameters, for example, tube current, slice thickness, and convolution kernel, and covered a wide range of imaging protocols for abdominal CT imaging. Moreover, our study was based on heterogeneous ROIs, that is, ROIs containing both object and background, which is an advantage over previous studies using homogenous ROI phantoms, for example, paper/rubber-filled cartridges (19).

Overall, for the ACR CT phantom, tube current affected reliability the most, slice thickness the second, and convolution kernel the least. The small effect of convolution kernels was due to the similarity of the 2 "smooth" kernels used in this abdominal study. The histogram-based RFs showed much higher reliability than textural RFs.

For tube current, 200 Effective mAs represented high-dose CT imaging, while 25 Effective mAs represented low-dose CT imaging. It is quite intuitive that CT images derived from high-dose scanning would yield more reliable RFs as it produced higher quality images than low noise scanning. Therefore, to obtain high RF reliability, high-dose CT imaging is recommended, especially for those radiomic studies using textural RFs. When keeping all other imaging acquisition parameters unchanged, increasing the slice thickness from 1.25 mm to 5 mm can reduce image noise by 50%. It is reasonable to believe that thick-section CT imaging yielded more reliable RFs. However, thick-section CT imaging introduces larger partial volume effect than thin-section CT imaging. In clinical practice, partial volume effect is one of the main negative effects that lowered image resolution and thus blurred fine structures within/around lesions, for example, small vessels, boundary of tumor margin, etc. It will also affect some RFs extracted from thick-section CT images. Therefore, the selection of RFs and slice thickness should depend on the aim of the radiomic study. Because the 2 convolution kernels, STANDARD and SOFT, both belonged to smooth soft-tissue kernels which yielded low-noise image, their influence on RF reliability was similar. Also, our results showed that smooth soft-tissue kernels used by abdominal CT scans had little impact on RF reliability.

In this study, 2 categories of RFs, histogram-based and the textural, were investigated. Histogram-based RFs showed much higher reliability than textural RFs, especially the *mean* and *std*. It is actually one of the basic requirements for a CT scanner that *mean* should be reliable across different image acquisition settings. Our results showed this. For the *std*, its high reliability was somewhat due to the use of the polyethylene and acrylic objects to create image patterns, which possessed dozens of Hounsfield unit (HU) intensity different from the water-equivalent background. Nevertheless, according to this finding, it is quite reliable to apply *std* in charactering tumor lesions with dozens of HU difference from the background, such as liver metastasis of colorectal cancer [mean, 68 HU; range, 40–115 HU as reported in the CRYSTAL clinical trials (25, 26)] and gastrointestinal stromal tumors [mean, 72 HU; range, 46–156 HU as reported in the Choi criteria study (27)].

In contrast to histogram-based RFs, more attention should be paid to the use of textural RFs. Textural RFs are easily affected by tube current, which is an imaging parameter directly proportional to patient radiation dose. High tube current guarantees high reliability of textural RFs, but leads to high patient dose. Therefore, the use of textural RFs should depend on the aim of a study. For example, it is inadvisable to use textural RFs in a low-dose CT screening study (28), whereas it might be safe to use textural RFs in a CT-based radiation therapy study (29).

There were several limitations of our pilot study. First, the created image patterns were simple, involving only 2 materials for each pattern, polyethylene, and a water-equivalent background, or acrylic and a water-equivalent background. Second, only a small set of RFs from 2 feature categories were investigated. Third, only 1 CT scanner was used. To address these limitations, we propose future studies, including designing more sophisticate phantoms that mimic in vivo lesions with the help of 3D-printing technique (30), using a high-throughput analysis method to evaluate a large scale of RFs (20), and involving multiple scanners from multiple institutions to attain CT images under more image acquisition settings (19).

CONCLUSION

In this study, we explored the reliability of RFs on multiple CT image acquisition settings. To the best of our knowledge, this is the first study investigating RF reliability by comparing real feature values calculated from scanned phantom images and reference feature values computed from computer-generated phantom images. We found that CT image acquisition settings influenced RF reliability to varying degrees. Therefore, attention should be paid when using RFs for CT-based radiomic studies, especially textural RFs.

Supplemental Materials

Supplemental Material: http://dx.doi.org/10.18383/j.tom.2016.00005.sup.01

ACKNOWLEDGMENTS

This work was supported in part by Grant U01 CA140207 and U01 CA225431 from the National Cancer Institute (NCI). The content is solely the responsibility of the authors and does not necessarily represent the funding sources.

Disclosures: No disclosures to report.

Conflict of Interest: The authors have no conflict of interest to declare.

REFERENCES

1. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR, Dekker A, Quackenbush J, Gillies RJ, Lambin P. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2015;278:563–577.
3. Gevaert O, Xu J, Hoang CD, Leung AN, Xu Y, Quon A, Rubin DL, Napel S, Plevritis SK. Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data–methods and preliminary results. Radiology. 2012;264:387–396.
4. Parmar C, Leijenaar RT, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, Rietbergen MM, Haibe-Kains B, Lambin P, Aerts HJ. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. Sci Rep. 2015;5:11044.
5. Liu Y, Kim J, Balagurunathan Y, Li Q, Garcia AL, Stringfield O, Ye Z, Gillies RJ. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas. Clin Lung Cancer. 2016;17:441–448.e6.
6. Zhou M, Scott J, Chaudhury B, Hall L, Goldgof D, Yeom KW, Iv M, Ou Y, Kalpathy-Cramer J, Napel S, Gillies R, Gevaert O, Gatenby R. Radiomics in brain tumor: image assessment, quantitative feature descriptors, and machine-learning approaches. AJNR Am J Neuroradiol. 2018;39:208–216.
7. Valdora F, Houssami N, Rossi F, Calabrese M, Tagliafico AS. Rapid review: radiomics and breast cancer. Breast Cancer Res Treat. 2018;169:217–229.
8. West DL, Kotrotsou A, Niekamp AS, Idris T, Giniebra Camejo D, Mazal NJ, Cardenas NJ, Goldberg JL, Colen RR. CT-based radiomic analysis of hepatocellular carcinoma patients to predict key genomic information [abstract]. JCO. 2017;35.15_suppl.e15623.
9. Jeong WK, Jamshidi N, Felker ER, Raman SS, Lu DS. Radiomics and radiogenomics of primary liver cancers. Clin Mol Hepatol. 2018. [Epub ahead of print].
10. Starmans MP, Miclea RL, van der Voort SR, Niessen WJ, Thomeer MG, Klein S, editors. Classification of malignant and benign liver tumors using a radiomics approach. Proc. SPIE. 2018; 105741D.
11. Yu H, Scalera J, Khalid M, Touret A-S, Bloch N, Li B, Qureshi MM, Soto JA, Anderson SW. Texture analysis as a radiomic marker for differentiating renal tumors. Abdom Radiol (NY). 2017;42:2470–2478.
12. van Rossum PS, Xu C, Fried DV, Goense L, Lin SH. The emerging field of radiomics in esophageal cancer: current evidence and future potential. Transl Cancer Res. 2016;5:410–23.
13. Cook GJ, Yip C, Siddique M, Goh V, Chicklore S, Roy A, Marsden P, Ahmad S, Landau D. Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? qj J Nucl Med. 2013;54:19–26.
14. Aerts HJ, Grossmann P, Tan Y, Oxnard GR, Rizvi N, Schwartz LH, Zhao B. Defining a radiomic response phenotype: a pilot study using targeted therapy in NSCLC. Sci Rep. 2016;6:33860.
15. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. Eur Radiol. 2012;22:796–802.
16. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldgof DB, Hall LO, Lambin P, Balagurunathan Y, Gatenby RA, Gillies RJ. Radiomics: the process and the challenges. Magn Reson Imaging. 2012;30:1234–1248.
17. Zhao B, Tan Y, Tsai WY, Qi J, Xie C, Lu L, Schwartz LH. Reproducibility of radiomics for deciphering tumor phenotype with imaging. Sci Rep. 2016;6:23428.
18. Lu L, Ehmke RC, Schwartz LH, Zhao B. Assessing agreement between radiomic features computed for multiple CT imaging settings. PloS One. 2016;11:e0166550.
19. Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, Rodriguez-Rivera E, Dodge C, Jones AK, Court L. Measuring CT scanner variability of radiomics features. Invest Radiol. 2015;50:757–765.
20. Berenguer R, Pastor-Juan MdR, Canales-Vázquez J, Castro-García M, Villas MV, Legorburo FM, Sabater S. Radiomics of CT features may be nonreproducible and redundant: influence of CT acquisition parameters. Radiology. 2018;288(2):407–415.
21. Li Q, Gavrielides MA, Sahiner B, Myers KJ, Zeng R, Petrick N. Statistical analysis of lung nodule volume measurements with CT in a large-scale phantom study. Med Phys. 2015;42:3932–3947.
22. Gavrielides MA, Berman BP, Supanich M, Schultz K, Li Q, Petrick N, Zeng R, Siegelman J. Quantitative assessment of nonsolid pulmonary nodule volume with computed tomography in a phantom study. Quant Imaging Med Surg. 2017;7:623–635.
23. McCollough CH, Bruesewitz MR, McNitt-Gray MF, Bush K, Ruckdeschel T, Payne JT, Brink JA, Zeman RK; American College of Radiology. The phantom portion of the American College of Radiology (ACR) computed tomography (CT) accreditation program: practical tips, artifact examples, and pitfalls to avoid. Med Phys. 2004;31:2423–2442.
24. Haralick RM, Shanmugam K. Textural features for image classification. IEEE Trans Syst Man Cybern. 1973:610–621.
25. Dercle L, Lu L, Lichtenstein P, Yang H, Wang D, Zhu J, et al. Impact of variability in portal venous phase acquisition timing in tumor density measurement and treatment response assessment: metastatic colorectal cancer as a paradigm. JCO Clin Cancer Inform. 2017;1:1–8.
26. Van Cutsem E, Köhne CH, Hitre E, Zaluski J, Chang Chien CR, Makhson A, D'Haens G, Pintér T, Lim R, Bodoky G, Roh JK, Folprecht G, Ruff P, Stroh C, Tejpar S, Schlichting M, Nippgen J, Rougier P. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. 2009;360:1408–1417.
27. Choi H, Charnsangavej C, Faria SC, Macapinlac HA, Burgess MA, Patel SR, Chen LL, Podoloff DA, Benjamin RS. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. J Clin Oncol. 2007;25:1753–1759.
28. National Lung Screening Trial Research Team, Church TR, Black WC, Aberle DR, Berg CD, Clingan KL, Duan F, Fagerstrom RM, Gareen IF, Gierada DS, Jones GC, Mahon I, Marcus PM, Sicks JD, Jain A, Baum S. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365:395–409.
29. Huynh E, Coroller TP, Narayan V, Agrawal V, Hou Y, Romano J, Franco I, Mak RH, Aerts HJ. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. Radiother Oncol. 2016;120:258–266.
30. Filippou V, Tsoumpas C. Recent advances on the development of phantoms using 3D printing for imaging with CT, MRI, PET, SPECT, and ultrasound. Med Phys. 2018;45:e740–e60.