Review of Value of CT Texture Analysis and Machine Learning in Differentiating Fat-Poor Renal Angiomyolipoma from Renal Cell Carcinoma

Yuhan Zhang[1], Xu Li[1], Yang Lv[2], and Xinquan Gu[1]

[1]Department of Urology, China-Japan Union Hospital of Jilin University, Changchun, China; and [2]Department of Anesthesia, The Second Hospital of Jilin University, Changchun, China

Corresponding Author:
Xinquan Gu, PhD
Department of Urology, China-Japan Union Hospital of Jilin University, 18F, Section 7, No. 126, Xiantai Street, Erdao District, Changchun, 130033, China;
E-mail: guxq@jlu.edu.cn

ABSTRACT

The diagnosis of patients with suspected angiomyolipoma relies on the detection of abundant macroscopic intralesional fat, which is always of no use to differentiate fat-poor angiomyolipoma (fp-AML) from renal cell carcinoma and diagnosis of fp-AML excessively depends on individual experience. Texture analysis was proven to be a potentially useful biomarker for distinguishing between benign and malignant tumors because of its capability of providing objective and quantitative assessment of lesions by analyzing features that are not visible to the human eye. This review aimed to summarize the literature on the use of texture analysis to diagnose patients with fat-poor angiomyolipoma vs those with renal cell carcinoma and to evaluate its current application, limitations, and future challenges in order to avoid unnecessary surgical resection.

INTRODUCTION

Renal angiomyolipoma (AML) containing smooth muscle cells, dysmorphic blood vessels, and adipose tissues (1, 2), is the most common benign solid renal tumor observed in clinical practice (3), most of which can be easily diagnosed by means of conventional computed tomography (CT) and magnetic resonance imaging (MRI) that can detect abundant macroscopic intralesional fat (4–6). Approximately 5% of renal AMLs have too little fat to be detected with either CT or MRI, making it difficult to differentiate fat-poor angiomyolipoma (fp-AML) from renal cell carcinoma (RCC) (7–9), a subtype of AML termed fp-AML or AML without visible fat. Fp-AMLs are the most frequent benign renal masses subjected to unnecessary surgery (10, 11). A variety of methods were proposed to differentiate fp-AML from RCC, such as angular interface, high attenuation of lesions at unenhanced CT, and strongly prolonged enhancement (12–14). However, these imaging findings showed insufficient specificity, inconsistent reproducibility, or inadequate prospective reliability (3, 15). The differentiation between benign and malignant tumors is of essential importance for the decision of proper treatment, but the diagnosis of fp-AML is challenging, time-consuming, and dependent much on the experience of individual radiologists.

As a branch of radiomics, quantitative texture analysis is an emerging technology that extracts and evaluates features from digital images, detects subtle changes and heterogeneity beyond human vision, and provides an objective method by analyzing the intensity, distribution, and relationship of pixel gray levels within a digital image (16). As an objective assessment of lesions, texture analysis assesses tumor heterogeneity and may reflect information about tissue characteristics (17, 18); this method was proven by a number of studies to be a potentially useful biomarker for the diagnosis, therapeutic response, and prognosis of colorectal, lung, esophageal, hepatic, and head and neck cancers (19–24). In recent years, several studies focusing on differentiating fp-AML from RCC have provided new approaches with high accuracy, sensitivity, and specificity by using CT texture analysis (CTTA) and machine learning as noninvasive methods. However, clinical urologists remain unfamiliar with the value of quantitative CTTA and machine learning in differentiating fp-AML from RCC, although CTTA is a promising biomarker. In this paper, the basic concept of texture analysis, its workflow, and application in the differentiation between fp-AML and RCC were provided, and we will discuss its current challenges and future development. This

Figure 1. Basic workflow of texture analysis.
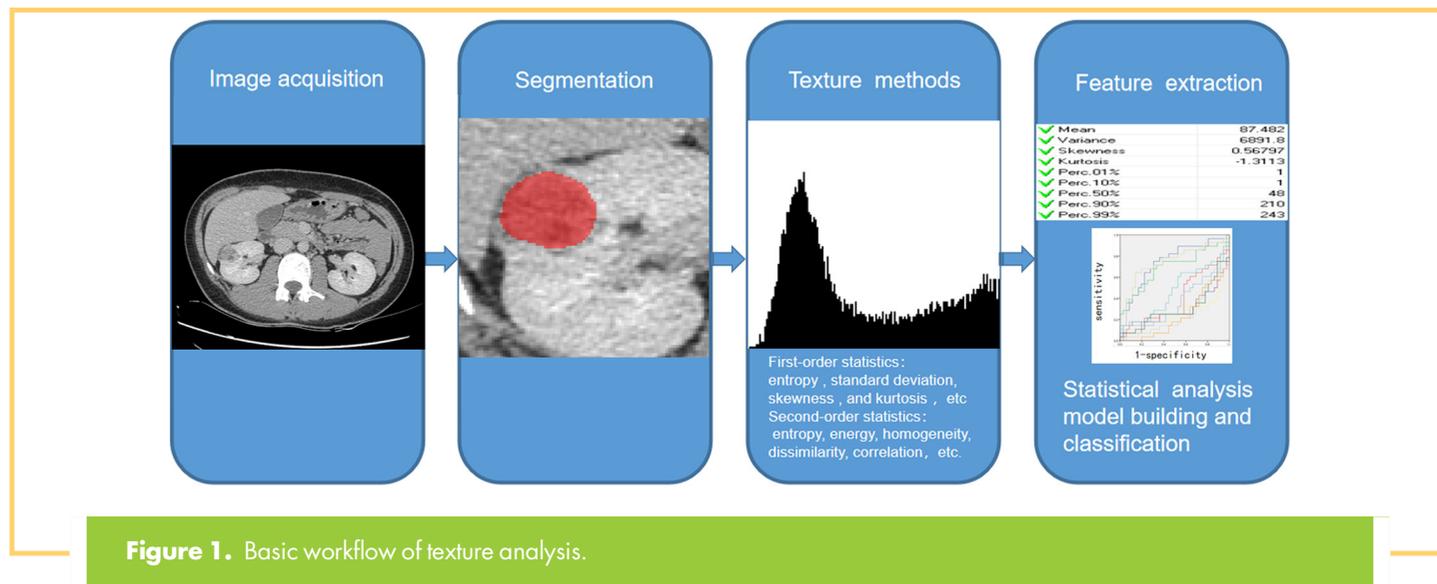
new technique may be beneficial to avoiding unnecessary surgical resection.

TEXTURE ANALYSIS: BASIC CONCEPT AND METHODOLOGY

In general, the basic workflow of CTTA includes image acquisition, segmentation, feature extraction, feature selection, statistical analysis, and classification (Figure 1).

CT Technique and Image Acquisition

Images of patient can be collected by the computerized search of picture archiving and communication systems. Patient characteristics are listed in Table 1, and study characteristics are listed in Table 2. No limits are set on CT image acquisition protocols. The standardization of protocols across medical imaging centers is typically lacking, which, however, is not a problem in the conventional identification of radiologic features used in clinical practice (25). Digital Imaging and Communications in Medicine format images were more popular for storage and analysis by articles that we were interested in, and most of the software available can handle this format of data. CT is the first-line imaging examination for the characterization of renal masses because of its good sensitivity and specificity (26, 27). As a result, CTTA is more convenient for texture analysis of renal masses in clinical practice. It is worth noting that images from different scanners may increase variability in the values of features calculated from CT images, and consideration should be given to striking a balance between the sufficient number of patients and data homogeneity (28).

Tumor Segmentation and Feature Extraction

Segmentation is of critical importance for images to be analyzed because subsequent feature data are generated from the region of interest (ROI) segmented from surrounding tissues (25). Various kinds of open-source software were developed to compute ROIs. Although this is a time-consuming work and a semi-automated approach was proven to be a quick method that can reduce the interobserver variability (29), most studies chose manual segmentation delineated by experienced radiologists. Notably, ROIs should be drawn at a distance of 2–3 mm from the tumor margin to minimize the partial volume effects of para-tumor renal parenchyma and perinephric fat (3, 30, 31). To be specific, some studies delineated lesions on enhanced CT images and applied and adjusted them in other phases to acquire the accurate ROI of each phase or took enhanced CT images as a reference (5, 32).

The core of radiomics is to extract feature data to quantitatively describe the attributes of ROIs. A variety of software packages, commercial and open-source, are available for researchers to extract features from delineated images. MaZda (3, 5), Pyradiomics (30, 32), in-house software MATLAB (33, 34), IBEX (35, 36), and TexRAD (37–39) were used to evaluate quantitative texture parameters in ROIs. Statistical-, model-, and transform-based methods were used for texture analysis; among these, statistical-based ones are the most commonly used to describe the relationship of gray-level values within an image (40).

This kind of features extracted from images is subdivided into 3 types, namely, first-, second-, and higher-order features. Specifically, first-order features evaluate the gray-level distribution from the pixel intensity histogram in an ROI, including mean intensity, skewness entropy, uniformity, threshold, kurtosis, and standard deviation (16). Second-order features focus on the image pattern of the spatial relationship or cooccurrence of pixel values in the ROI, including entropy, contrast, energy, and homogeneity. Gray-level cooccurrence matrix and gray-level run-length matrix are the 2 most common methods (40). Aiming to analyze the relationship between pixels ($\geq$3), higher-order features are less analyzed and used in studies.

Feature Selection, Statistical Analysis, Modeling, and Classification

Features extracted from ROIs may be large in number, and they will not contribute equally and are not even relevant to differentiate fp-AML from RCC. Feature selection is of essential importance

Table 1. Patient Characteristics

Study	Patients	No. of Masses	No. of Fp-AML	Renal Cell Carcinoma				Age (Year)		Tumor Size (mm)	
				ccRCC	pRCC	chRCC	Others	AML	RCC	AML	RCC
Hodgdon et al. (5)	100	100	16	51	13	20	0	53 ± 12	59 ± 13	18 ± 13	24 ± 9
Takahashi et al. (49)	153	172	24	98	36	14a		53 ± 14	60 ± 12	15 ± 7	21 ± 8
Feng et al. (31)	58	58	17	31	2	6	2	48.7 ± 10.8	56.2 ± 12.3	28 ± 9	32 ± 7
Cui et al. (30)	168	171	41	82	22	26	0	48.56 ± 12.90	55.27 ± 11.56 (cc) 49.27 ± 12.99 (p) 55.00 ± 11.80 (ch)	<4 0	<4 0
You et al. (33)	67	67	17	50	0	0	0	47.53 ± 2.76	53.32 ± 1.62	21.06 ± 11.32	24.66 ± 1.14
Deng et al. (51)	377	385	31	249	49	56	0	NM	59 ± 13	NM	45 ± 35
Varghese et al. (50)	147	147	18	85	23	21	0	NM	NM	NM	NM
Yan et al. (3)	48	50	18	18	14	0	0	44.5, range 26–61	53.9, range 36–79 (cc) 57.6, range34-77 (p)	28.47 ,range 8–51	33.22; range, 15–49 (cc) 33.09, range14–51 (p)
Yang G. et al. (53)	58	58	32	0	0	24	0	50.38 + 8.66	52.88 + 10.86	NM	NM
Yang R. et al. (32)	163	163	45	95	10	13	0	48.6 ± 13.7	52.9 ± 13.1	25, range 21–33	29, range 24–33

Abbreviations: AML, angiomyolipoma; RCC, renal cell carcinoma; cc, clear cell carcinoma; p, papillary renal carcinoma; Ch, chromophobe renal carcinoma; NM, not mentioned.
[a]Includes chromophobe renal carcinoma and other RCCs.

to select optimal features and avoid overtraining with a poor outcome. Generally, the number of texture features calculated from images is much larger than the sample size of patients. Hence, the reduction of dimensionality may be important to reduce the risk of type I errors and overfitting (41).

Different from traditional statistical methods, machine learning classifiers are performed to process data. Machine learning can be defined to enable computers to make predictions based on past experience. As a branch of artificial intelligence, it has advanced rapidly in the past decade with the development of computational resource. In some fields other than medicine, such as natural language processing and traffic volume prognosis, machine learning plays a central role. Various kinds of difficult tasks such as diagnosis, prognosis, and response of therapy have been solved using this new technology (42–47). This objective technique has no subjective disadvantage and can process tremendous medical data. Although it helps radiologists deal with complex problems, it has caught few urologists' attention. The final goal of machine learning in these studies is to obtain an effective diagnostic model including multiple relevant parameters with high accuracy to differentiate fp-AML from RCCs. Support vector machine (SVM) is the most common method in these studies, along with logistic regression (LR), k-nearest neighbors, and random forest. Despite the various methods, the performance of machine learning was always evaluated by receiver operating characteristic curve and accuracy in clinical tasks (48).

DISCUSSION
The results of univariate and multivariate analyses are listed in Table 3.

Univariate Analysis
The univariate analysis was performed with traditional statistical method, for both first- and second-order features. Despite the variability in texture analysis, entropy showed promising results in differentiating fp-AML from RCC (5, 31, 33, 49–51). Hodgdon et al. conducted research on unenhanced CT images and claimed that RCC can be characterized by a higher level of entropy than fp-AML ($P \leq .01$) (5). Similar result was reported by You et al., who found a higher degree of entropy dissimilarity and a lower degree of energy and homogeneity in clear cell RCC in the corticomedullary phase (33). Deng et al. observed that entropy >5.62 had a high specificity of 85.7% for predicting RCC but has a sensitivity of 31.3% (51). Previous studies suggested that higher lesion entropy was a strong predictor of RCC, and greater entropy was consistently observed in RCC compared with fp-AML. Entropy measured the complexity or disorder of images and represented the heterogeneity of tumors (31). In addition to entropy, RCC was labeled with a higher degree of dissimilarity and a lower level of lesion homogeneity.

Multivariate Analysis
Multivariate analysis was performed to accurately differentiate RCC from fp-AML. Machine learning, the core of artificial

Table 2. Study Characteristics

Study	Publication Year	Study Period	Institution	Pathology Method	Processors and Readers		
					NO.	Experience (year)	Blinded
Hodgdon et al. (5)	2015	January 2002–August 2013	The Ottwa hospital	Surgical resection	2	3/8	NM
Takahashi et al. (49)	2015	January 2003–January 2011	Mayo clinic	Surgical resection	1	13	Yes
Feng et al. (31)	2017	June 2013–September 2016	The third Xiangya hospital	Surgical resection	2	7/8	NM
Cui et al. (30)	2019	January 2008–September 2017	Jiangmen central hospital	Surgical resection	2	NM	NM
You et al. (33)	2019	November 2008–December 2010	Asian medical center	NM	1	6	NM
Varghese et al. (50)	2018	June 2009–June 2015	University of Southern California	Surgical resection	1	NM	NM
Yan et al. (3)	2015	January 2008–April 2014	Guangdong general hospital	Biopsy or surgical resection	2	16/36	NM
Yang G. et al. (53)	2019	June 2009–January 2018	The affiliated hospital of Qingdao university	Surgical resection	2	8/20	NM
Deng et al. (51)	2019	October 2005–October 2016	Mayo clinic	NM	1	15	Yes
Yang R. et al. (32)	2019	January 2012–December 2018	Guangzhou first people's hospital	NM	2	3/14	Yes

Abbreviation: NM = not mentioned.

intelligence, is widely applied to achieve better outcomes and more accurate diagnostic ability, whose goal is to obtain a classifier or a model with high accuracy. LR is a popular classifier in these mathematical models because of its simplicity and popularity with researchers (52). Hodgdon et al. established an LR model by combining the top 3 texture features per session, resulting in an AUC of 0.89 ± 0.04, which was significantly >0.5. The sensitivity and specificity of identifying RCC ranged from 87% to 93% and from 63% to 75%, respectively (5). Yang et al. applied least absolute shrinkage and selection operator LR to develop a 2D texture model (AUC, 0.811; 95% CI, 0.695–0.927) and a 3D texture model (AUC, 0.915; 95% CI, 0.838–0.993), which showed good discrimination and calibration in distinguishing fp-AML from clear cell RCC (53). Takahashi et al. built a LR model with entropy, demographic data, shape features, and subjective heterogeneity factors and differentiated small fp-AML from RCC with a sensitivity and specificity of 50% and 98%, respectively (49). In addition to LR, SVM was widely used in the studies included in this paper. Feng et al. developed an SVM classifier with 11 features selected by the SVM-RFE (SVM with the recursive feature elimination) method and achieved the highest accuracy, sensitivity, specificity, and AUC of 93.9%, 87.8%, 100%, and 0.955, respectively, in differentiating fp-AML from RCC (31). Lee et al. showed that the model comprising relief feature selection and SVM classifier achieved an accuracy, sensitivity, specificity, and AUC of 72.1% ± 4.2%, 71.0% ± 5.1%, 73.2% ± 6.1%, and 0.717 ± 0.045, respectively (54). In addition, k-nearest neighbors, random forest, and nonlinear discriminant analysis were applied and proven to be reliable methods of differentiating fp-AML from RCC (3, 54).

Both unenhanced and enhanced CT images were incorporated into these studies. Hodgdon et al. restricted their study to the analysis of unenhanced CT images for the reason that little literature focused on the effect of iodinated contrast material on texture analysis (5). Textural differences extracted from unenhanced CT images were independent of contrast effects. It happened that a similar case appeared in the article of Cui et al. who found that unenhanced images performed the best in differentiating fp-AML from RCC in single-phase texture analysis and made significant contributions in the 3-phase group (30). In addition, some studies (31–33) found significant differences in univariate analysis during the unenhanced phase ($P < .05$). Furthermore, the models based on these unenhanced CT images can significantly decrease radiation exposure and benefit patients suffering from renal insufficiency.

Tumor heterogeneity, which is difficult to quantify with traditional imaging methods, was proven to be greater in malignant tumors than benign ones (55). Although the heterogeneity of renal mass is a crucial feature to differentiate RCC from fp-AML (7), the subjective analysis of heterogeneity depends too much on experience of readers and lacks reproducibility. Recent studies suggested that the objective quantification of heterogeneity evaluated by the methods of standard deviation, entropy, and uniformity was of help to differentiate AML from RCC (56), which was consistent with the results of univariate analysis included in this review. Hodgdon et al. claimed that lower lesion homogeneity and higher lesion entropy were biomarkers of RCCs (5). Yang et al. reported the 3 top-ranked texture features extracted from ROIs, and 2 of them were gray-level nonuniformity and size zone

Table 3. Methods, Results, and Performance

Study	Phases	Segmentation	Extraction	Machine Learning	Discriminative Features	Best Performance of Models			
						SEN	SPE	ACC	AUC
Hodgdon et al. (5)	UN	Manually	MaZda, version 4.6	SVM LR	Mean gray-level, angular second moment, gray-level entropy, sum entropy, and sum average	88% (LR)	75% (LR)	83%–91% (SVM)	0.89 ±0.04 (LR)
Takahashi et al. (49)	UN CE-CT	NM	Matlab (MathWorks)	LR	Entropy	50%	98%	NM	0.943
Feng et al. (31)	UN CMP NP	Manually	CT kinetics (version1.20, GEHealthcare)	SVM	Skewness, mean, median, 10th, 25th, 75th, and 90th percentiles (UP), energy and entropy (UN, CMP, and NP)	87.8%	100%	93.9%	0.955
Cui et al. (30)	UN CMP NP	Manually	PyRadiomics (version3.6.5)	SVM	NM	89.23%	96.15%	92.69%	0.96
You et al. (33)	UN CMP NP EP	Manually	Matlab (MathWorks)	SVM	Mean (UN), SD, homogeneity, dissimilarity, energy, and entropy (CMP)	82%	76%	85%	0.85
Varghese et al. (50)	UN CMP NP EP	Manually	Matlab (MathWorks)	LR	NM	NM	NM	NM	0.95–0.98
Yan et al. (3)	UN CMP NP	Manually	MaZda, version 4.6	kNN artificial neural classifer	NM	NM	NM	90.7%-100%	NM
Yang G. et al. (53)	CMP NP EP	Manually	Radiomics cloud platform V2.1.2	LASSO	NM	93.75%	79.17%	87.5%	0.915
Deng et al. (51)	Portal venous phase	Manually	TexRAD, version 3.9	LR	Entropy, maximum positive pixel	33%	97%	NM	0.658
Yang R. et al. (32)	UN CMP NP EP	Manually	PyRadiomics	SVM, LR, Random forest, Bagging	90th percentile, mean, median, root mean squared, skewness, IMC1, IMC2, GLN, and SZN	0.83	0.82	0.82	0.90

Abbpreviations: NM, not mentioned; UN, unenhanced; CE, contrast-enhanced; CMP, corticomedullary phase; NP, nephrographic phase; EP, excreory phase; SD, standard deviation; LR, logistic regression; SVM, support vector machine; kNN, k-nearest neighbor; LASSO, least absolute shrinkage and selection operator; SEN, sensitivity; SPE, specificity; ACC, accuracy; AUC, area under curve; IMC1, informational measure of correlation 1 of the GLCM texture feature; GLN, gray-level nonuniformity of the GLSZM texture feature; SZN, size zone nonuniformity of the GLSZM texture feature.

nonuniformity, which were markers of tumor homogeneity, showing that fp-AML was more homogenous than the RCC. Similar results were obtained in other articles (31, 33, 50). Tumor heterogeneity is a feature of malignancy, and a lesion with increased heterogeneity is likely related to tumor angiogenesis, cellular infiltration, and areas of necrosis (5, 55). Histological evidence that the inner components of fp-AML appear to be more regular than those of RCC in terms of cell proliferation and less-invasive potential supports these findings (32).

LIMITATIONS

Intralesional fat on CT or MRI is the typical characteristic of AML. However, some tumors contain too little fat to be detected, which makes it difficult to differentiate them from RCCs (57). In the past decade, texture analysis, an emerging brunch of radiomics, has shown promising potential to distinguish between malignant and benign tumors. This kind of new technology has rapidly developed with the increasing digitalization in the hospital and progress in image acquisition protocols, along with easier access to the picture archiving and communication systems. It is an objective approach and automatic extraction of quantitative features from images, which differs from traditional radiology methods depending too much on the subjective visual interpretation and expertise of radiologists and urologists (58). In addition, texture analysis is capable of helping with the diagnosis of both common and rare tumors and even differentiating benign and malignant lymph nodes in patients with primary lung cancer (58–63). In the past 5 years, focus was put on this area to differentiate fp-AML from RCC. The articles included in this review used multiple feature extraction and classification methods, and all achieved relatively satisfactory results (AUC > 0.5).

Despite having a foreseeable optimistic and promising future, texture analysis and machine learning encounter problems to be solved in clinical decision-making (25, 64). The limitation of clinical implementation and use mostly results from the lack of standards and reproducibility. Regarded as one of the

foundations for scientific research, reproducibility plays an important role. Nonreproducible consequences waste the time and money of researchers (65). The variety of CT scanners, different methods of delineating ROIs (manually, semiautomatically, automatically), and the inhomogeneity of software (commercial, open-source, or developed in-house) used to extract and process features may be responsible for this issue. Every step in the workflow of texture analysis should be standardized to achieve a better and more convincing outcome. Recently, there have been efforts to standardize the definitions and flow, and studies have been conducted on the reliability and stability of features to enhance the reproducibility (66–69).

Another limitation is the relatively small sample size. Without universal standardized workflows, large centralized data repositories, or image data-sharing methods, researchers always fight their own battle, which may give rise to limited data. Besides, type I error and overfitting may be unavoidable owing to the limited size of samples. It is suggested that statistical corrections such as Holm–Bonferroni sequential correction should be applied and sample size should be 5–10 times of texture features analyzed to reduce these problems (40, 70).

Most studies included were retrospective studies (case–control studies), which were known sometimes to overestimate the sensitivity and specificity of diagnosis and lead to biases (71). Severe biases can produce adverse consequences, such as results with errors and incorrect conclusions. Hence, a well-designed and prospective study should be conducted to clarify the results achieved by the articles focusing on differentiating fp-AML from RCC. Besides, most studies on texture analysis showed only the correlation between features and results, which, however, did not mean causation (72).

SUMMARY
Despite the disadvantages that we discussed in the previous section, we can initially give the conclusion that CTTA can be useful for the differentiation of fp-AML from RCC on both unenhanced CT and enhanced CT. Texture features such as entropy that showed promising potential may be regarded as quantitative, noninvasive, and effective imaging biomarkers. Models made by machine learning–based methods performed with open-source software or algorithms with high accuracy are encouraging for the future imaging studies. However, deficiency and limitation exist, and universally accepted standards need to be established. Before the implementation into widespread clinical practice, this kind of new technology requires further validation on a larger scale.

REFERENCES
1. Katabathina VS, Vikram R, Nagar AM, Tamboli P, Menias CO, Prasad SR. Mesenchymal neoplasms of the kidney in adults: imaging spectrum with radiologic-pathologic correlation. Radiographics. 2010;30:1525–1540.
2. Steiner MS, Goldman SM, Fishman EK, Marshall FF. The natural history of renal angiomyolipoma. J Urol. 1993;150:1782–1786.
3. Yan L, Liu Z, Wang G, Huang Y, Liu Y, Yu Y, Liang C. Angiomyolipoma with minimal fat: differentiation from clear cell renal cell carcinoma and papillary renal cell carcinoma by texture analysis on CT images. Acad Radiol. 2015;22:1115–1121.
4. Israel GM, Hindman N, Hecht E, Krinsky G. The use of opposed-phase chemical shift MRI in the diagnosis of renal angiomyolipomas. AJR Am J Roentgenol. 2005;184:1868–1872.
5. Hodgdon T, McInnes MD, Schieda N, Flood TA, Lamb L, Thornhill RE. Can quantitative CT texture analysis be used to differentiate fat-poor renal angiomyolipoma from renal cell carcinoma on unenhanced CT images?. Radiology. 2015;276:787–796.
6. Woo S, Suh CH, Cho JY, Kim SY, Kim SH. Diagnostic performance of CT for diagnosis of fat-poor angiomyolipoma in patients with renal masses: a systematic review and meta-analysis. AJR Am J Roentgenol. 2017;209:W297–W307.
7. Jinzaki M, Tanimoto A, Narimatsu Y, Ohkuma K, Kurata T, Shinmoto H, Hiramatsu K, Mukai M, Murai M. Angiomyolipoma: imaging findings in lesions with minimal fat. Radiology. 1997;205:497–502.
8. Kim JY, Kim JK, Kim N, Cho KS. CT histogram analysis: differentiation of angiomyolipoma without visible fat from renal cell carcinoma at CT imaging. Radiology. 2008;246:472–479.
9. Hindman N, Ngo L, Genega EM, Melamed J, Wei J, Braza JM, Rofsky NM, Pedrosa I. Angiomyolipoma with minimal fat: Can it be differentiated from clear cell renal cell carcinoma by using standard MR techniques? Radiology. 2012;265:468–477.
10. Fujii Y, Komai Y, Saito K, Iimura Y, Yonese J, Kawakami S, Ishikawa Y, Kumagai J, Kihara K, Fukui I. Incidence of benign pathologic lesions at partial nephrectomy for presumed RCC renal masses: Japanese dual-center experience with 176 consecutive patients. Urology. 2008;72:598–602.
11. Kutikov A, Fossett LK, Ramchandani P, Tomaszewski JE, Siegelman ES, Banner MP, Van Arsdalen KN, Wein AJ, Malkowicz SB. Incidence of benign pathologic findings at partial nephrectomy for solitary renal mass presumed to be renal cell carcinoma on preoperative imaging. Urology. 2006;68:737–740.
12. Zhang YY, Luo S, Liu Y, Xu RT. Angiomyolipoma with minimal fat: differentiation from papillary renal cell carcinoma by helical CT. Clin Radiol. 2013;68:365–370.
13. Kim JK, Park SY, Shon JH, Cho KS. Angiomyolipoma with minimal fat: differentiation from renal cell carcinoma at biphasic helical CT. Radiology. 2004;230:677–684.
14. Kim YH, Han K, Oh YT, Jung DC, Cho NH, Park SY. Morphologic analysis with computed tomography may help differentiate fat-poor angiomyolipoma from renal cell carcinoma: a retrospective study with 602 patients. Abdom Radiol (NY). 2018;43:647–654.
15. Yang C-W, Shen S-H, Chang Y-H, Chung H-J, Wang J-H, Lin AT, Chen K-K. Are there useful CT features to differentiate renal cell carcinoma from lipid-poor renal angiomyolipoma? AJR Am J Roentgenol. 2013;201:1017–1028.
16. Soni N, Priya S, Bathla G. Texture analysis in cerebral gliomas: a review of the literature. AJNR Am J Neuroradiol. 2019;40:928–934.
17. Lubner MG, Stabo N, Lubner SJ, del Rio AM, Song C, Halberg RB, Pickhardt PJ. CT textural analysis of hepatic metastatic colorectal cancer: pre-treatment tumor heterogeneity correlates with pathology and clinical outcomes. Abdom Imaging. 2015;40:2331–2337.
18. Yasaka K, Akai H, Mackin D, Court L, Moros E, Ohtomo K, Kiryu S. Precision of quantitative computed tomography texture analysis using image filtering: a phantom study for scanner variability. Medicine (Baltimore). 2017;96:e6993.
19. Ng F, Ganeshan B, Kozarski R, Miles KA, Goh V. Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. Radiology. 2013;266:177–184.
20. 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.
21. Yip C, Landau D, Kozarski R, Ganeshan B, Thomas R, Michaelidou A, Goh V. Primary esophageal cancer: heterogeneity as potential prognostic biomarker in patients treated with definitive chemotherapy and radiation therapy. Radiology. 2014;270:141–148.

22. Zhang H, Graham CM, Elci O, Griswold ME, Zhang X, Khan MA, Pitman K, Caudell JJ, Hamilton RD, Ganeshan B, Smith AD. Locally advanced squamous cell carcinoma of the head and neck: CT texture and histogram analysis allow independent prediction of overall survival in patients treated with induction chemotherapy. Radiology. 2013;269:801–809.

23. Huang YL, Chen JH, Shen WC. Diagnosis of hepatic tumors with texture analysis in nonenhanced computed tomography images. Acad Radiol. 2006;13:713–720.

24. Cui C, Cai H, Liu L, Li L, Tian H, Li L. Quantitative analysis and prediction of regional lymph node status in rectal cancer based on computed tomography imaging. Eur Radiol. 2011;21:2318–2325.

25. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278:563–577.

26. Davenport MS, Neville AM, Ellis JH, Cohan RH, Chaudhry HS, Leder RA. Diagnosis of renal angiomyolipoma with Hounsfield unit thresholds: effect of size of region of interest and nephrographic phase imaging. Radiology. 2011;260:158–165.

27. Krishna S, Murray CA, McInnes MD, Chatelain R, Siddaiah M, Al-Dandan O, Narayanasamy S, Schieda N. CT imaging of solid renal masses: pitfalls and solutions. Clin Radiol. 2017;72:708–721.

28. Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, Rodriguez-Rivera E, Dodge C, Jones AK, Court L. Measuring computed tomography scanner variability of radiomics features. Invest Radiol. 2015;50:757–765.

29. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, Mitra S, Shankar BU, Kikinis R, Haibe-Kains B, Lambin P, Aerts HJWL. Robust radiomics feature quantification using semiautomatic volumetric segmentation. PloS One. 2014;9: e102107.

30. Cui E-M, Lin F, Li Q, Li R-G, Chen X-M, Liu Z-S, Long W-S. Differentiation of renal angiomyolipoma without visible fat from renal cell carcinoma by machine learning based on whole-tumor computed tomography texture features. Acta Radiol. 2019;60:1543–1552.

31. Feng Z, Rong P, Cao P, Zhou Q, Zhu W, Yan Z, Liu Q, Wang W. Machine learning-based quantitative texture analysis of CT images of small renal masses: differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol. 2018;28:1625–1633.

32. Yang R, Wu J, Sun L, Lai S, Xu Y, Liu X, Ma Y, Zhen X. Radiomics of small renal masses on multiphasic CT: accuracy of machine learning-based classification models for the differentiation of renal cell carcinoma and angiomyolipoma without visible fat. Eur Radiol. 2020;30:1254–1263.

33. You MW, Kim N, Choi HJ. The value of quantitative CT texture analysis in differentiation of angiomyolipoma without visible fat from clear cell renal cell carcinoma on four-phase contrast-enhanced CT images. Clin Radiol. 2019;74:547–554.

34. Fried DV, Tucker SL, Zhou S, Liao Z, Mawlawi O, Ibbott G, Court LE. Prognostic value and reproducibility of pretreatment CT texture features in stage III non-small cell lung cancer. Int J Radiat Oncol Biol Phys. 2014;90:834–842.

35. Fried DV, Mawlawi O, Zhang L, Fave X, Zhou S, Ibbott G, Liao Z, Court LE. Stage III non-small cell lung cancer: prognostic value of FDG PET quantitative imaging features combined with clinical prognostic factors. Radiology. 2016;278:214–222.

36. van Rossum PSN, Fried DV, Zhang L, Hofstetter WL, van Vulpen M, Meijer GJ, Court LE, Lin SH. The incremental value of subjective and quantitative assessment of 18F-FDG PET for the prediction of pathologic complete response to preoperative chemoradiotherapy in esophageal cancer. J Nucl Med. 2016;57:691–700.

37. Ganeshan B, Strukowska O, Skogen K, Young R, Chatwin C, Miles K. Heterogeneity of focal breast lesions and surrounding tissue assessed by mammographic texture analysis: preliminary evidence of an association with tumor invasion and estrogen receptor status. Front Oncol. 2011;1:33.

38. Yip C, Davnall F, Kozarski R, Landau DB, Cook GJR, Ross P, Mason R, Goh V. Assessment of changes in tumor heterogeneity following neoadjuvant chemotherapy in primary esophageal cancer. Dis Esophagus. 2015;28:172–179.

39. Raman SP, Schroeder JL, Huang P, Chen Y, Coquia SF, Kawamoto S, Fishman EK. Preliminary data using computed tomography texture analysis for the classification of hypervascular liver lesions: generation of a predictive model on the basis of quantitative spatial frequency measurements—a work in progress. J Comput Assist Tomogr. 2015;39:383–395.

40. Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT texture analysis: definitions, applications, biologic correlates, and challenges. Radiographics. Inc2017;37:1483–1503.

41. Larue RT, Defraene G, De Ruysscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. Br J Radiol. 2017;90:20160665.

42. Giger ML. Machine Learning in Medical Imaging. J Am Coll Radiol. 2018;15:512–520.

43. Ali HR, Dariush A, Provenzano E, Bardwell H, Abraham JE, Iddawela M, Vallier A-L, Hiller L, Dunn JA, Bowden SJ, Hickish T, McAdam K, Houston S, Irwin MJ, Pharoah PDP, Brenton JD, Walton NA, Earl HM, Caldas C. Computational pathology of pretreatment biopsies identifies lymphocyte density as a predictor of response to neoadjuvant chemotherapy in breast cancer. Breast Cancer Res. 2016;18:21.

44. Huang C-H, Zeng C, Wang Y-C, Peng H-Y, Lin C-S, Chang C-J, Yang H-Y. A study of diagnostic accuracy using a chemical sensor array and a machine learning technique to detect lung cancer. Sensors (Basel). 2018;18.

45. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH, Cannon TD, Krystal JH, Corlett PR. Cross-trial prediction of treatment outcome in depression: a machine learning approach. Lancet Psychiatry. 2016;3:243–250.

46. Cha KH, Hadjiiski L, Chan H-P, Weizer AZ, Alva A, Cohan RH, Caoili EM, Paramagul C, Samala RK. Bladder cancer treatment response assessment in CT using radiomics with deep-learning. Sci Rep. 2017;7:8738.

47. Garapati SS, Hadjiiski L, Cha KH, Chan H-P, Caoili EM, Cohan RH, Weizer A, Alva A, Paramagul C, Wei J, Zhou C. Urinary bladder cancer staging in CT urography using machine learning. Med Phys. 2017;44:5814–5823.

48. Petrick N, Sahiner B, Armato SG, Bert A, Correale L, Delsanto S, Freedman MT, Fryd D, Gur D, Hadjiiski L, Huo Z, Jiang Y, Morra L, Paquerault S, Raykar V, Samuelson F, Summers RM, Tourassi G, Yoshida H, Zheng B, Zhou C, Chan H-P. Evaluation of computer-aided detection and diagnosis systems. Med Phys. 2013;40 087001.

49. Takahashi N, Leng S, Kitajima K, Gomez-Cardona D, Thapa P, Carter RE, et al. Small (< 4 cm) renal masses: differentiation of angiomyolipoma without visible fat from renal cell carcinoma using unenhanced and contrast-enhanced CT. AJR Am J Roentgenol. 2015;205:1194–1202.

50. Varghese BA, Chen F, Hwang DH, Cen SY, Desai B, Gill IS, Duddalwar VA. Differentiation of predominantly solid enhancing lipid-poor renal cell masses by use of contrast-enhanced CT: Evaluating the role of texture in tumor subtyping. AJR Am J Roentgenol. 2018;211:W288–w96.

51. Deng Y, Soule E, Cui E, Samuel A, Shah S, Lall C, Sundaram C, Sandrasegaran K. Usefulness of CT texture analysis in differentiating benign and malignant renal tumors. Clin Radiol. 2020;75:108–115.

52. Avanzo M, Stancanello J, El Naqa I. Beyond imaging: the promise of radiomics. Phys Med. 2017;38:122–139.

53. Yang G, Gong A, Nie P, Yan L, Miao W, Zhao Y, Wu J, Cui J, Jia Y, Wang Z. Contrast-enhanced CT texture analysis for distinguishing fat-poor renal angiomyolipoma from chromophobe renal cell carcinoma. Mol Imaging. 2019;18:1536012119883161.

54. Lee HS, Hong H, Jung DC, Park S, Kim J. Differentiation of fat-poor angiomyolipoma from clear cell renal cell carcinoma in contrast-enhanced MDCT images using quantitative feature classification. Med Phys. 2017;44:3604–3614.

55. Davnall F, Yip CSP, Ljungqvist G, Selmi M, Ng F, Sanghera B, Ganeshan B, Miles KA, Cook GJ, Goh V. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? Insights Imaging. 2012;3:573–589.

56. Leng S, Takahashi N, Gomez Cardona D, Kitajima K, McCollough B, Li Z, Kawashima A, Leibovich BC, McCollough CH. Subjective and objective heterogeneity scores for differentiating small renal masses using contrast-enhanced CT. Abdom Radiol (NY). 2017;42:1485–1492.

57. Flum AS, Hamoui N, Said MA, Yang XJ, Casalino DD, McGuire BB, Perry KT, Nadler RB. Update on the Diagnosis and Management of Renal Angiomyolipoma. J Urol. 2016;195:834–846.

58. Kassner A, Thornhill RE. Texture analysis: a review of neurologic MR imaging applications. AJNR Am J Neuroradiol. 2010;31:809–816.

59. Schieda N, Thornhill RE, Al-Subhi M, McInnes MDF, Shabana WM, van der Pol CB, Flood TA. Diagnosis of sarcomatoid renal cell carcinoma with CT: evaluation by qualitative imaging features and texture analysis. AJR Am J Roentgenol. 2015;204:1013–1023.

60. Li Z, Yu L, Wang X, Yu H, Gao Y, Ren Y, Wang G, Zhou X. Diagnostic performance of mammographic texture analysis in the differential diagnosis of benign and malignant breast tumors. Clin Breast Cancer. 2018;18:e621–e7.

61. Wibmer A, Hricak H, Gondo T, Matsumoto K, Veeraraghavan H, Fehr D, Zheng J, Goldman D, Moskowitz C, Fine SW, Reuter VE, Eastham J, Sala E, Vargas HA. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. Eur Radiol. 2015;25:2840–2850.

62. Digumarthy SR, Padole AM, Lo Gullo R, Singh R, Shepard JO, Kalra MK. CT texture analysis of histologically proven benign and malignant lung lesions. Medicine (Baltimore). 2018;97:e11172.

63. Bayanati H, E. Thornhill R, Souza CA, Sethi-Virmani V, Gupta A, Maziak D, Amjadi K, Dennie C. Quantitative CT texture and shape analysis: can it differentiate benign and malignant mediastinal lymph nodes in patients with primary lung cancer? Eur Radiol. 2015;25:480–487.

64. Summers RM. Texture analysis in radiology: does the emperor have no clothes? Abdom Radiol (NY). 2017;42:342–345.

65. Eisner DA. Reproducibility of science: fraud, impact factors and carelessness. J Mol Cell Cardiol. 2018;114:364–368.

66. Jaggi A, Mattonen SA, McNitt-Gray M, Napel S. Stanford DRO Toolkit: digital reference objects for standardization of radiomic features. Tomography. 2020;6:111–117.

67. Lu L, Liang Y, Schwartz LH, Zhao B. Reliability of radiomic features across multiple abdominal CT image acquisition settings: a pilot study using ACR CT phantom. Tomography. 2019;5:226–231.

68. McNitt-Gray M, Napel S, Jaggi A, Mattonen SA, Hadjiiski L, Muzi M, et al. Standardization in quantitative imaging: a multicenter comparison of radiomic features from different software packages on digital reference objects and patient data sets. Tomography. 2020;6:118–128.

69. Paul R, Shafiq U, Hassan M, Moros EG, Gillies RJ, Hall LO, Goldgof DB. Deep feature stability analysis using CT images of a physical phantom across scanner manufacturers, cartridges, pixel sizes, and slice thickness. Tomography. 2020;6:250–260.

70. Patel N, Henry A, Scarsbrook A. The value of MR textural analysis in prostate cancer. Clin Radiol. 2019;74:876–885.

71. Whiting PF, Rutjes AW, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. J Clin Epidemiol. 2013;66:1093–1104.

72. Napel S, Giger M. Special section guest editorial: radiomics and imaging genomics: quantitative imaging for precision medicine. J Med Imaging (Bellingham). 2015;2 041001.