



Machine learning analysis of adrenal lesions: preliminary study evaluating texture analysis in the differentiation of adrenal lesions

Canan Altay 
İşıl Başara Akın 
Abdullah Hakan Özgül 
Süleyman Cem Adıyaman 
Abdullah Serkan Yener 
Mustafa Seçil 

PURPOSE

This study aimed to determine the accuracy of texture analysis in differentiating adrenal lesions on unenhanced computed tomography (CT) images.

METHODS

In this single-center retrospective study, 166 adrenal lesions in 140 patients (64 women, 76 men; mean age 56.58 ± 13.65 years) were evaluated between January 2015 and December 2019. The lesions consisted of 54 lipid-rich adrenal adenomas, 37 lipid-poor adrenal adenomas (LPAs), 56 adrenal metastases (ADM), and 19 adrenal pheochromocytomas (APhs). Each adrenal lesion was segmented by manually contouring the borders of the lesion on unenhanced CT images. A texture analysis of the CT images was performed using Local Image Feature Extraction software. First-order and second-order texture parameters were assessed, and 45 features were extracted from each lesion. One-Way analysis of variance with Bonferroni correction and the Mann-Whitney U test was performed to determine the relationships between the texture features and adrenal lesions. Receiver operating characteristic curves were performed for lesion discrimination based on the texture features. Logistic regression analysis was used to generate logistic models, including only the texture parameters with a high-class separation capacity (i.e., $P < 0.050$). SPSS software was used for all statistical analyses.

RESULTS

First-order and second-order texture parameters were identified as significant factors capable of differentiating among the four lesion types ($P < 0.050$). The logistic models were evaluated to ascertain the relationships between LPA and ADM, LPA and APh, and ADM and APh. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy of the first model (LPA vs. ADM) were 85.7%, 70.3%, 81.3%, 76.4%, and 79.5%, respectively. The sensitivity, specificity, PPV, NPV, and accuracy of the second model (LPA vs. APh) were all 100%. The sensitivity, specificity, PPV, NPV, and accuracy of the third model (ADM vs. APh) were 87.5%, 82%, 36.8%, 98.2%, and 82.7%, respectively.

CONCLUSION

Texture features may help in the characterization of adrenal lesions on unenhanced CT images.

KEYWORDS

Adrenal adenoma, adrenal glands, adrenal mass, computed tomography, texture analysis

From the Department of Radiology (C.A. ✉ drcananaltay@gmail.com, I.B.A., A.H.Ö., M.S.), Dokuz Eylül University Faculty of Medicine, Izmir, Turkey; Department of Endocrinology (S.C.A., A.S.Y.), Dokuz Eylül University Faculty of Medicine, Izmir, Turkey.

Received; 04 April 2021, revision requested; 02 May 2021, last revision received; 05 December 2021, accepted 10 January 2022.



Epub: 06.02.2023

Publication date: 29.03.2023

DOI: 10.5152/dir.2022.21266

Incidental adrenal masses are common lesions, with a rate of detection of 4%–6% using abdominal computed tomography (CT) in patients undergoing abdominal imaging.¹ Adrenal adenomas are the most common adrenal lesions, with a prevalence up to 9% in the general population.² The rate of adrenal adenomas increases with age.³ Nearly 25% of adrenal adenomas have insufficient intracytoplasmic lipid content for accurate detection using conventional CT, rendering these lesions difficult to differentiate from malignant adrenal lesions and

You may cite this article as: Altay C, Başara Akın I, Özgül AH, Adıyaman SC, Yener AS, Seçil M. Machine learning analysis of adrenal lesions: preliminary study evaluating texture analysis in the differentiation of adrenal lesions. *Diagn Interv Radiol.* 2023;29(2):234-243.

adrenal pheochromocytomas (APhs).³ This can ultimately lead to unnecessary surgical resection of these benign adrenal lesions. Clinical findings and radiological features in unenhanced CT images have a limited ability to differentiate lipid-poor adrenal adenomas (LPAs) from malignant adrenal lesions and APhs because they are in the same density spectrum range. CT washout imaging and magnetic resonance imaging (MRI) are non-invasive imaging techniques used for the differential diagnosis of adrenal lesions.⁴⁻⁶

Lipid-rich adrenal adenomas (LRAs) can be easily recognized in unenhanced CT, with homogeneous and relatively low attenuation values [<10 Hounsfield units (HU)], due to their high fat content.⁷ LRAs can also be diagnosed through MRI with high sensitivity and specificity.¹ Most LRAs show a loss of signal intensity ($>20\%$) in out-of-phase images from conventional dual-phase T1-weighted imaging and using the Dixon technique due to intratumoral microscopic fat content.⁴ LPAs, however, may have a diagnostic overlap with other lipid-poor adrenal lesions, such as adrenal metastasis (ADM) and APhs, in unenhanced CT or MRI images.⁸ Washout CT of the adrenal gland is the most reliable imaging method for differentiating adrenal adenomas from other lesions. Most adrenal adenomas show a rapid washout pattern after administration of the contrast agent.⁴ However, hypervascular ADM, especially from hepatocellular carcinoma or renal cell carcinoma, may also show a rapid washout pattern.⁹ Therefore, the definitive preoperative diagnosis of adrenal lesions is not always possible.

Texture analysis (TA), performed using several complex mathematical processes, is an objective assessment of the structure, gray-level intensity, and pixel position of a le-

sion.¹⁰ The TA parameters of the radiological images can be divided into first-order and second-order statistics. First-order TA parameters consist of the uniformity (in gray-level dispersion), skewness (histogram asymmetry), entropy (irregularity in gray-level dispersion), kurtosis (histogram flatness), energy (uniformity of the distribution), mean/minimum/maximum density, and standard deviation of the gray-level histogram dispersion.¹¹ Second-order TA, used to assess the spatial relationships among pixels, includes the following parameters: gray-level co-occurrence matrix (GLCM), neighborhood gray-level difference matrix (NGLDM), gray-level run-length matrix (GLRLM), gray-level zone length matrix (GLZLM), and their subgroup parameters.¹² First-order and second-order TA parameters in adrenal lesions may help physicians to differentiate subgroup lesions based on the underlying histopathological composition. This study investigated the role of TA in differentiating adrenal lesions in unenhanced CT images.

Materials and Methods

Patient selection

The protocol of this retrospective study was approved by our institutional review board (approval number: 2020/20-31; date: 2020-08-31), and written informed consent was obtained from each patient. The endocrinology department database was reviewed to identify patients with LRA, LPA, ADM, or APh between January 2015 and December 2019. The inclusion criteria were age ≥ 19 years, histopathologically proven APh, diagnosis of LRA, LPA, or ADM based on radiological features, positron emission tomography (PET)-CT features, long-term follow-up, histopathological diagnosis, or biopsy, and unenhanced CT performed before resection or during follow-up screening. Eligible patients had an adrenal lesion determined by unenhanced CT or PET-CT to be ADM (for those with known primary tumors with standardized uptake value (SUV)_{max} ratio >2.5 [the accepted value for distinguishing between benign and cancerous lesions], and SUV_{max} values of ADM were higher than those of normal liver parenchyma), LRA (≤ 10 HU in unenhanced CT images), or LPA (diagnosis made by histopathological or radiological follow-up at least 36 months following initial CT).¹³

Histopathological diagnosis of adrenal lesions is made by surgical intervention, and adrenal lesion biopsy should only be considered for making a diagnosis of metastatic

disease in patients with known or suspected non-adrenal cancer.¹⁴ The exclusion criteria were adrenal lesion <1 cm ($n = 18$), lack of unenhanced CT images ($n = 82$), CT images with severe motion or other artifacts ($n = 7$), and CT imaging performed at another institution ($n = 23$). Both lesions of patients with bilateral adrenal lesions were included in the study. We were unable to confirm a diagnosis of adenoma versus ADM histopathologically in three patients, and these patients were also excluded. After these criteria had been applied, a total of 140 patients with 166 adrenal lesions were included (Figure 1). All patients were evaluated through non-contrast abdominal CT using the CT devices available in our department (Supplementary Table 1).

Endocrinological examinations were performed in all cases preoperatively. Metanephrine and normetanephrine were measured in 24-h urine for a differential diagnosis of pheochromocytoma in all patients. Basal plasma levels of adrenocorticotropic hormone and serum levels of cortisol, the dexamethasone suppression test, plasma renin activity, and the aldosterone/renin ratio were used to evaluate adrenal adenoma.

Texture analysis

The unenhanced CT images of each patient were reviewed by three radiologists (HAÖ, IBA, and CA) with 4, 15, and 20 years of experience in abdominal radiology, respectively. The selected images of all adrenal lesions were decided by consensus. For each adrenal lesion, a selected axial image was obtained through CT for use in TA. The selected images for the 166 adrenal masses were anonymized and exported from Sectra IDS7 PACS (Sectra AB, Linköping, Sweden). The image selection was made in the largest diameter of the lesion. Reprocessing of the images was performed using standardization: (gray value-average value)/standard deviation. The images were examined and segmented using Local Image Feature Extraction (LIFEX) software (version 5.1; <http://www.lifexsoft.org>) independently by two radiologists (IBA and CA) with 15 and 20 years of experience in abdominal radiology, respectively. Regions of interest (ROIs) were drawn manually on each CT image for the measurements (Figures 2, 3). To provide uniformity, the number of gray levels was adjusted to 128 (7 bits). Furthermore, to homogenize the voxel sizes in the XYZ directions, they were adjusted as X, 0.5 mm; Y, 0.5 mm; and Z, 2.5 mm after calculating their mean ± 3 standard deviations. First-order and second-order TA features were calculated in each adrenal lesion.

Main points

- Adrenal masses are commonly encountered in daily practice, and it can be difficult to distinguish between lipid-poor adenoma and masses that may require surgical treatment based on non-contrast computed tomography (CT).
- Texture analysis reveals minor differences in gray levels, which can provide guidance in the differential diagnosis of visceral neoplasms.
- This study suggests that lipid-poor adenoma, adrenal metastasis, and pheochromocytoma can be distinguished through an analysis of textural features in non-contrast CT images

Statistical Analysis

IBM SPSS software for Windows, version 24.0 (IBM Corporation, Armonk, NY), was used for all statistical analyses. We recorded demographic data, including the patient's age, sex, diagnosis, and TA measurements. Interobserver variability between the two radiologists was evaluated using intraclass correlation coefficients (0–0.4, poor agreement; 0.41–0.6, moderate agreement; 0.61–0.8, good agreement; 0.81–1, excellent agreement). We performed One-Way analysis of variance (ANOVA) with Bonferroni correction to compare the differences in the TA parameters of LRA, LPA, ADM, and APh. The non-parametric Mann–Whitney U test was performed to detect significant associations between the texture parameters and each adrenal subgroup. LPAs were excluded from the statistical evaluation because they are easily recognized in non-contrast CT examinations and do not cause diagnostic difficulties. Receiver operating characteristic (ROC) curve analysis was performed to evaluate the

diagnostic performance of the parameters in differentiating among the adrenal lesion subgroups by calculating the area under the ROC curve (AUC). A *P* value <0.050 was considered statistically significant.

Next, we developed three models to identify texture features predicting LPA, ADM, and APh, with *P* < 0.050 in the univariate analysis. A logistic regression model was developed based on the qualitative features. In the final logistic regression model, statistically significant texture parameters were used for the three comparisons (12 for LPA vs. ADM, 15 for LPA vs. APh, and 4 for ADM vs. APh). The model performance was assessed according to sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), and accuracy.

Results

The final study population consisted of 140 patients (64 men, 76 women; mean age, 56.58 ± 13.65 years) with 166 adrenal lesions.

Fifty-four adrenal lesions were diagnosed with LRA, 37 were diagnosed with LPA, 56 were diagnosed with ADM, and 19 were diagnosed with APh (Table 1). There were no significant differences in age or sex among the subgroups. In our series of 91 adrenal adenomas, 76% were non-functioning. The total interobserver agreement between two radiologists (IBA and CA) was $\kappa = 0.73$ in the TA calculation. Due to the high level of interobserver agreement and the large amount of numerical data, the data of observer 1 are presented in the tables. Twenty-seven texture parameters, consisting of 11 histogram features, 7 GLCM features, 3 GLRLM features, 3 NGLDM features, and 3 GLZLM features, showed statistically significant differences. The One-Way ANOVA identified 16 second-order texture parameters that were significantly different among all adrenal lesion subgroups (Table 2). The Mann–Whitney U test identified 12 texture parameters that could differentiate between LPA and ADM (*P* = 0.000–0.032), 15 that

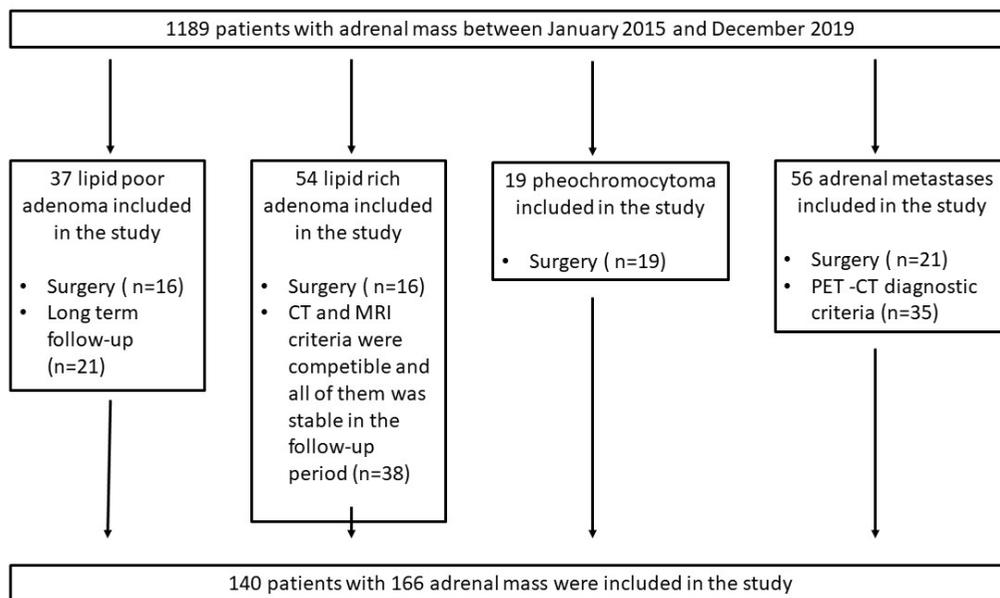


Figure 1. Flowchart of patients with adrenal lesions. CT, computed tomography; MRI, magnetic resonance imaging; PET-CT, positron emission tomography-computed tomography.

Table 1. Imaging characteristics of adrenal lesions

Imaging characteristics	Lipid-rich adenoma (n = 54)	Lipid-poor adenoma (n = 37)	Metastasis (n = 56)	Pheochromocytoma (n = 19)
HU _{min}	-70.57 ± 22.9	-57.12 ± 32.7	-32.90 ± 30.2	-30.67 ± 34.7
HU _{mean}	-9.84 ± 13.9	14.60 ± 14.6	30.32 ± 13.9	33.47 ± 11.9
HU _{max}	49.56 ± 141.2	74 ± 98.9	83.5 ± 26.4	89 ± 19.4
Mean lesion size (cm)	5.4	5.9	7.8	4.9
Mean SUV _{max}	-	-	8.6	-

SUV_{max}, maximum standardized uptake value; HU_{min}, Hounsfield unit minimum; HU_{mean}, Hounsfield unit mean; HU_{max}, Hounsfield unit maximum.

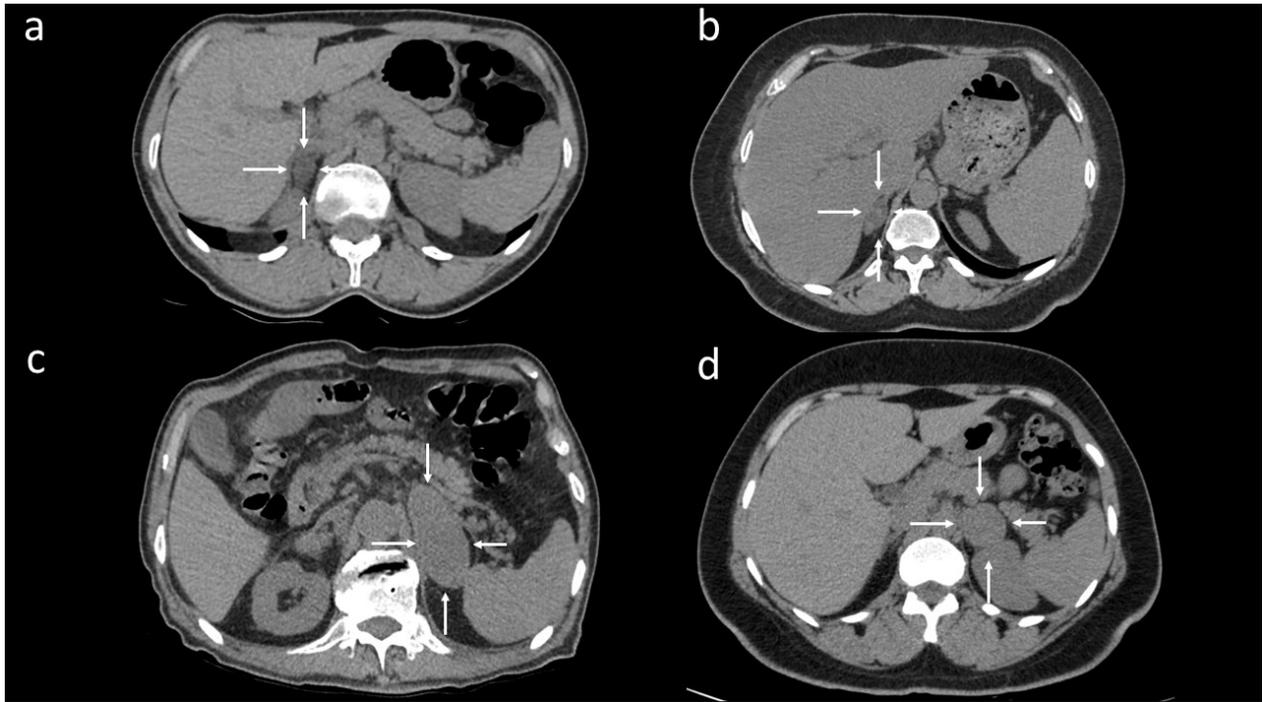


Figure 2. (a) Lipid-rich adenoma in a 49-year-old woman. A well-circumscribed and low-density lesion was observed in the right adrenal gland in axial computed tomography (CT) (white arrows). (b) A 37-year-old woman with lipid-poor adenoma diagnosed histopathologically. Axial unenhanced CT showed a lipid-poor right adrenal mass (white arrows). The mass with a density of 22–56 Hounsfield unit (HU) had a heterogeneous internal structure. (c) Adrenal metastases were diagnosed through surgery in a 56-year-old man with lung cancer. Axial unenhanced CT showed a large, well-defined, homogenous left adrenal mass with a density of 54 HU (white arrows). (d) Surgically resected pheochromocytoma in a 32-year-old man. Axial unenhanced CT showed a well-defined, homogenous, spherical left adrenal mass with a density of 46 HU (white arrows).

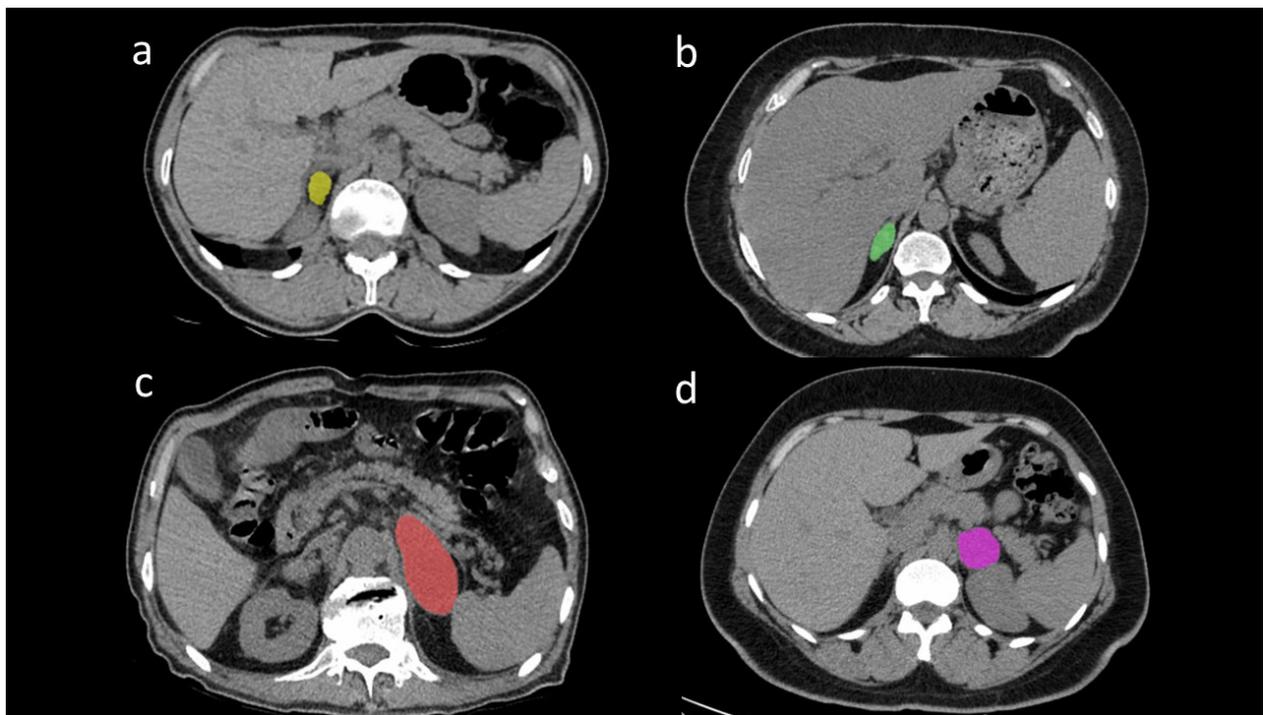


Figure 3. Segmentation examples of reader 1 (a) and reader 2 (b). (a) Texture analyses of the lipid-rich adrenal adenoma using local image feature extraction (LIFEx) software [lesion tagged with a yellow region of interest (ROI)]. (b) Texture analyses of the lipid-poor adrenal adenoma using LIFEx software (lesion tagged with a green ROI). (c) Texture analyses of the adrenal metastasis using LIFEx software (lesion tagged with a red ROI). (d) Texture analyses of the adrenal pheochromocytoma using LIFEx software (lesion tagged with a pink ROI).

Table 2. Differences in texture parameters for differentiating adrenal lesions on non-enhanced computed tomography

Texture analysis parameters	1-lipid-rich adenoma (n = 54)			2-lipid-poor adenoma (n = 37)			3-metastasis (n = 56)			4-pheochromocytoma (n = 19)			P value
	Median	Q1 (IQR 25%)	Q3 (IQR 75%)	Median	Q1 (IQR 25%)	Q3 (IQR 75%)	Median	Q1 (IQR 25%)	Q3 (IQR 75%)	Median	Q1 (IQR 25%)	Q3 (IQR 75%)	
HU _{min}	-70.57	-87.59	-54.44	-57.12	-79.71	-34.92	-32.90	-51.02	-22.63	-30.67	-52	-14	<0.001
HU _{mean}	-9.84	-18.05	-0.17	14.60	-0.51	24.98	30.32	23.50	35.13	33.47	28.16	38.91	<0.001
HU _{max}	49.56	35.86	60.29	74	62.72	92.03	83.5	68.59	107.50	89	74.8	98.6	0.015
HUQ1	-20.21	-30.71	-12.62	1.0	-12.29	14.12	18.30	10.73	22.20	21	16.53	28.14	<0.001
HUQ2	-10.03	-17.43	-0.34	14.57	-0.74	25.83	30.18	23.33	35.84	33	28.18	38.25	<0.001
HUQ3	2.91	-4.25	13.79	27.36	16.99	39.28	42.51	35.48	47.84	44.49	37.77	53.66	<0.001
HU-Skewness	-0.014	-0.012	-0.12	-0.065	-0.27	-0.025	-0.079	-0.21	-0.025	-0.101	-0.17	-0.022	0.068
HU-Kurtosis	3.15	2.96	3.53	3.16	2.92	3.63	3.21	3.03	3.57	3.20	3.04	3.61	0.896
HU-Excess Kurtosis	0.15	-0.034	0.53	0.167	-0.71	0.64	0.217	0.039	0.578	0.20	0.041	0.612	0.896
HU peak sphere 0.5 mL	0.46	0.46	0.46	0.470	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.010
HU peak sphere 1 mL	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.01	1.00	1.00	0.110
GLCM-homogeneity	0.1235	0.1186	0.1325	0.1287	0.1222	0.1361	0.1288	0.123	0.140	0.122	0.115	0.133	0.101
GLCM-energy	0.0011	0.0007	0.0016	0.0010	0.0005	0.0017	0.0008	0.0005	0.0012	0.0005	0.0004	0.0007	0.001
GLCM-contrast	473.94	378.78	532.39	412.12	351.81	489.62	429.78	356.13	494.84	503.14	408.77	580.48	0.058
GLCM-correlation	0.46	0.404	0.544	0.55	0.44	0.60	0.517	0.442	0.595	0.434	0.354	0.524	0.022
GLCM-entropy log10	2.97	2.79	3.17	2.99	2.79	3.28	3.14	2.92	3.38	3.30	3.17	3.43	<0.001
GLCM-entropy log2	9.90	9.29	10.54	9.94	9.29	10.90	10.45	9.70	11.22	10.98	10.53	11.41	<0.001
GLCM-dissimilarity	17.12	15.23	18.20	15.86	14.98	17.40	16.26	14.69	17.48	17.51	15.94	18.91	0.053
GLRLM-LRLGE	0.00304	0.0006	0.0048	0.0024	0.0009	0.0047	0.0058	0.0021	0.0083	0.0037	0.0018	0.0056	0.061
GLRLM-GLNU	9.67	6.49	14.89	8.69	5.36	17.44	13.85	7.64	28.70	25.84	14.90	34.45	<0.001
GLRLM-RLNU	638.61	384.49	1011.57	612.07	375.34	1340.15	936.18	490.26	1973.91	1630.25	1040.04	2418.51	<0.001
NGLDM-coarseness	0.0114	0.760	1.155	0.0106	0.0034	0.017	0.0091	0.0042	0.0137	0.0045	0.0029	0.0065	0.001
NGLDM-contrast	0.926	0.404	0.544	0.913	0.637	1.07	0.765	0.636	0.944	0.785	0.694	1.114	0.027
NGLDM-busyness	0.015	0.011	0.022	0.012	0.009	0.027	0.017	0.0126	0.0313	0.031	0.023	0.043	<0.001
GLZM-LZLGE	0.0032	0.0007	0.0071	0.0034	0.0010	0.0084	0.0063	0.0024	0.0146	0.0039	0.0025	0.0161	0.030
GLZM-GLNU	9.00	5.97	13.91	8.11	4.94	16.46	12.81	7.12	26.65	25.35	14.64	32.99	<0.001
GLZM-ZLNU	488.62	320.81	851.50	506.1	306.4	1163.5	763.88	414.42	1656.78	1321.87	874.36	2022.33	<0.001

GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; GLNU, gray-level non-uniformity; NGLDM, neighborhood gray-level difference matrix; HU_{min}, hounsfield unit minimum; HU_{mean}, hounsfield unit mean; HU_{max}, hounsfield unit maximum.

could differentiate between LPA and APh ($P = 0.000-0.039$), and 7 that could differentiate between ADM and APh ($P = 0.002-0.042$) (Table 3). Texture parameters that were significantly different between LPA and ADM,

between LPA and APh, and between ADM and APh were identified by ROC analysis. The AUCs for all independent factors distinguishing the lesions in these three comparisons were between 0.25 and 0.87. The AUCs, 95%

confidence intervals, optimal cut-off values, sensitivities, and specificities for each of the three comparisons are listed in Tables 3-5, respectively. Logistic regression analyses produced three logistic models. The logis-

Table 3. First-order and second-order radiomic features with significant differences between LRA, LPA, ADM, and APh

First- and second-order texture parameters	1-2*	1-3*	1-4*	2-3*	2-4*	3-4*
HU _{min}	0.032	<0.001	<0.001	0.001	0.007	0.443
HU _{mean}	<0.001	<0.001	<0.001	<0.001	<0.001	0.116
HU _{std}	0.021	0.995	0.138	0.032	0.003	0.144
HU _{max}	<0.001	<0.001	<0.001	0.027	0.036	0.692
HUQ1	<0.001	<0.001	<0.001	<0.001	<0.001	0.075
HUQ2	<0.001	<0.001	<0.001	<0.001	<0.001	0.144
HUQ3	<0.001	<0.001	<0.001	<0.001	<0.001	0.348
HU-skewness	0.067	0.032	0.028	0.100	0.869	0.635
HU peak sphere 0.5 mL	0.822	0.292	0.013	0.254	0.038	0.002
HU peak sphere 1 mL	0.181	0.306	0.033	0.624	0.206	0.100
GLCM-homogeneity	0.170	0.037	0.497	0.689	0.191	0.073
GLCM-energy	0.539	0.006	<0.001	0.101	0.011	0.095
GLCM-contrast	0.098	0.047	0.327	0.100	0.079	0.040
GLCM-correlation	0.025	0.031	0.346	0.510	0.033	0.053
GLCM-entropy log10	0.545	0.005	<0.001	0.103	0.009	0.093
GLCM-entropy log2	0.545	0.005	<0.001	0.103	0.009	0.093
GLCM-dissimilarity	0.106	0.035	0.365	0.826	0.085	0.040
GLRLM-LRLGE	0.904	0.021	0.315	0.031	0.222	0.407
GLRLM-GLNU	0.704	0.008	<0.001	0.023	0.002	0.042
GLRLM-RLNU	0.765	0.006	<0.001	0.089	0.005	0.062
NGLDM-coarseness	0.314	0.055	<0.001	0.666	0.039	0.008
NGLDM-contrast	0.450	0.002	0.191	0.124	0.697	0.355
NGLDM-busyness	0.359	0.061	<0.001	0.024	0.001	0.005
GLZM-LZLGE	0.790	0.009	0.163	0.031	0.121	0.526
GLZM-GLNU	0.710	0.008	<0.001	0.023	<0.001	0.007

(1) lipid-rich adrenal adenoma, (2) lipid-poor adrenal adenoma, (3) adrenal metastasis, (4) adrenal pheochromocytoma * $P < 0.050$. LRA, lipid-rich adrenal adenomas; LPA, lipid-poor adrenal adenoma, ADM, adrenal metastase; APh, adrenal pheochromocytoma; HU_{min}, hounsfield unit minimum; HU_{mean}, hounsfield unit mean; HU_{max}, hounsfield unit maximum; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; GLNU, gray-level non-uniformity; NGLDM, neighborhood gray-level difference matrix.

Table 4. Accuracy of radiomic features for differentiating between lipid-poor adrenal adenoma and adrenal metastasis

	AUC \pm std (95% confidence interval)	P value	Cut-off value	Sensitivity (%)	Specificity (%)
HU _{min}	0.71 \pm 0.054 (0.61–0.81)	0.012	–44.61	64.3	64.8
HU _{mean}	0.78 \pm 0.051 (0.69–0.88)	<0.001	24.00	73.2	72.9
HU _{std}	0.37 \pm 0.058 (0.25–0.48)	0.032	19.16	62.2	37.8
HU _{max}	0.63 \pm 0.060 (0.52–0.75)	0.027	80.32	58.9	59.4
HUQ1	0.78 \pm 0.052 (0.68–0.88)	<0.001	12.54	73.2	72.9
HUQ2	0.76 \pm 0.051 (0.68–0.88)	<0.001	23.86	73.2	72.9
HUQ3	0.63 \pm 0.052 (0.66–0.86)	<0.001	35.75	71.4	70.2
GLRLM-LRLGE	0.63 \pm 0.060 (0.51–0.75)	0.031	0.0036	58.9	59.4
GLRLM-GLNU	0.64 \pm 0.061 (0.52–0.76)	0.023	10.71	60.7	59.4
NGLDM-busyness	0.64 \pm 0.061 (0.52–0.76)	0.024	0.0124	76.8	45.9
GLZM-LZLGE	0.63 \pm 0.059 (0.52–0.75)	0.031	0.0022	76.8	59.5
GLZM-GLNU	0.64 \pm 0.060 (0.52–0.76)	0.023	6.75	76.8	62.2

HU_{min}, Hounsfield unit minimum; HU_{mean}, Hounsfield unit mean; HU_{max}, Hounsfield unit maximum; AUC, area under the ROC curve; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; GLNU, gray-level non-uniformity; NGLDM, neighborhood gray-level difference matrix; std, standard.

tic models were evaluated to ascertain the relationships between LPA and ADM, LPA and APh, and ADM and APh (Tables 4-6). The sensitivity, specificity, PPV, NPV, and accuracy of the first model (LPA vs. ADM) were 85.7%, 70.3%, 81.3%, 76.4%, and 79.5%, respectively ($P = 0.050$). The sensitivity, specificity, PPV,

NPV, and accuracy of the second model (LPA vs. APh) were all 100% ($P < 0.001$). The sensitivity, specificity, PPV, NPV, and accuracy of the third model (ADM vs. APh) were 87.5%, 82%, 36.8%, 98.2%, and 82.7%, respectively ($P = 0.018$). We present all values of our logistic regression models in Table 7. The first lo-

gistic model found that HU_{mean} , GLRLM-long run low gray-level emphasis (GLRLM-LGE), and NGLDM-Busyness were significant factors differentiating LPA from ADM, with odds ratios (ORs) of 3.87, 2.64, and 5.07, respectively. In the second model, GLRLM-gray-level non-uniformity (GLRLM-GLNU) and HU_{mean}

Table 5. Accuracy of radiomic features for differentiating between lipid-poor adrenal adenoma and adrenal pheochromocytoma

	AUC \pm std (95% confidence interval)	<i>P</i> value	Cut-off value	Sensitivity (%)	Specificity (%)
HU_{min}	0.72 \pm 0.079 (0.57–0.88)	0.071	–41.05	68.4	67.5
HU_{mean}	0.86 \pm 0.048 (0.77–0.96)	<0.001	26.96	78.9	78.4
HU_{std}	0.25 \pm 0.076 (0.10–0.40)	0.003	18.49	31.6	32.4
HU_{max}	0.67 \pm 0.071 (0.53–0.81)	0.040	82.26	63.2	62.1
HUQ1	0.87 \pm 0.048 (0.77–0.96)	<0.001	15.22	78.9	78.37
HUQ2	0.85 \pm 0.051 (0.75–0.95)	<0.001	24.42	78.9	78.4
HUQ3	0.83 \pm 0.054 (0.72–0.94)	<0.001	37.41	78.9	75.6
GLCM energy	0.29 \pm 0.072 (0.15–0.43)	0.013	0.00066	36.8	32.4
GLCM correlation	0.32 \pm 0.079 (0.17–0.48)	0.030	0.49	36.8	35.1
GLCM-entropy log 10	0.72 \pm 0.072 (0.58–0.86)	0.001	3.203	68.4	67.6
GLCM-entropy log 2	0.72 \pm 0.072 (0.58–0.86)	0.001	10.63	68.4	67.5
GLRLM-GLNU	0.76 \pm 0.070 (0.62–0.89)	0.002	15.19	73.7	72.9
GLRLM-RLNU	0.73 \pm 0.072 (0.59–0.87)	0.005	1046.48	73.7	70.2
NGLDM-coarseness	0.33 \pm 0.074 (0.18–0.47)	0.040	0.0045	47.4	32.4
NGLDM-busyness	0.78 \pm 0.065 (0.65–0.91)	0.001	0.022	78.9	72.9
GLZLM-GLNU	0.81 \pm 0.058 (0.69–0.92)	<0.001	14.29	78.6	73.9
GLZLM-ZLNU	0.73 \pm 0.072 (0.58–0.87)	0.006	896.77	73.7	70.3

HU_{min} , hounsfield unit minimum; HU_{mean} , hounsfield unit mean; HU_{max} , hounsfield unit maximum; AUC, area under the ROC curve; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; GLNU, gray-level non-uniformity; NGLDM, neighborhood gray-level difference matrix; std, standard.

Table 6. Accuracy of radiomic features for differentiating between adrenal metastasis and adrenal pheochromocytoma

	AUC \pm std (95% confidence interval)	<i>P</i> value	Cut-off value	Sensitivity (%)	Specificity (%)
HU peak sphere 0.5 mL	0.68 \pm 0.083 (0.51–0.84)	0.023	0.456	89.3	10.7
GLCM-contrast	0.66 \pm 0.079 (0.50–0.81)	0.040	462.22	63.2	62.5
GLCM-dissimilarity	0.66 \pm 0.078 (0.50–0.81)	0.040	16.85	63.2	62.5
GLRLM-GLNU	0.66 \pm 0.067 (0.52–0.79)	0.042	20.35	63.2	64.2
NGLDM-coarseness	0.29 \pm 0.065 (0.16–0.42)	0.008	0.0030	73.7	19.6
NGLDM-busyness	0.72 \pm 0.064 (0.59–0.84)	0.005	0.0249	68.4	64.2
GLZLM-GLNU	0.71 \pm 0.059 (0.59–0.82)	0.007	19.16	68.4	64.3

HU, hounsfield unit; AUC, area under the ROC curve; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; GLNU, gray-level non-uniformity; NGLDM, neighborhood gray-level difference matrix; std, standard.

Table 7. Statistical values of logistic regression models

	Model 1*	Model 2**	Model 3***
Sensitivity (%)	85.7	100	87.5
Specificity (%)	70.3	100	82
Positive predictive value (%)	81.3	100	36.8
Negative predictive value (%)	76.4	100	98.2
Accuracy (%)	79.5	100	82.7
<i>P</i> value	0.050	<0.001	0.018

*, statistical model between lipid-poor adrenal adenoma and adrenal metastasis; **, statistical model between lipid-poor adrenal adenoma and adrenal pheochromocytoma; ***, statistical model between adrenal metastasis and adrenal pheochromocytoma).

were identified as factors differentiating LPA from APH, with ORs of 119.09 and 4.98, respectively. An HU peak sphere of 0.5 mL was identified as a factor differentiating ADM from APH, with an OR of 4.2.

Discussion

We evaluated the performance of CT-based TA for differentiating adrenal lesions in unenhanced CT images. The results indicate that TA in unenhanced CT images may help to differentiate adrenal lesion subgroups. Numerous TA parameters can be useful for the differential diagnosis of adrenal lesions. GLCM-homogeneity, GLCM-contrast, GLRLM-LGE, and NGLDM-contrast are unique identifiers for discrimination, demonstrating the highest sensitivity and specificity rates (Table 1).

Unenhanced CT, CT washout, and adrenal MRI have been used to differentiate LRAs from other adrenal masses, with high sensitivities and specificities.^{3,4} However, there are occasionally false-positive diagnoses of LRAs that mimic ADM or APH.¹⁵ Adrenal LRAs are easily identified through contrast-enhanced CT using a cut-off of 10 HU.³ Washout CT is widely used to differentiate LRA and LPA from other adrenal lesions. Johnson et al.¹⁶ reported a sensitivity of 88% and specificity of 96% for diagnosing adrenal adenomas using absolute percentage washout in CT. However, Patel et al.¹⁷ revealed that LPA and APH have a diagnostic overlap with CT attenuation and washout criteria. Seo et al.¹⁸ compared the ability of MRI and washout CT to diagnose LPA and reported sensitivities of 75.7% and 100% and specificities of 60% and 80%, respectively. Hypervascular metastases in the adrenal gland show a rapid washout pattern in CT and cannot be differentiated from adrenal adenomas.¹⁹ Another diagnostic issue is that hepatocellular carcinomas and renal cell carcinomas are rich in intracytoplasmic fat, resulting in difficult differentiation from LRAs based on chemical-shift MRI.¹⁹ Therefore, the diagnostic capabilities of unenhanced CT, washout CT, and adrenal MRI are limited in terms of definitive differentiation among LPA, ADM, and APH.

TA enables the evaluation of a variety of image pixels used to describe the relationships among their gray-level intensity positions within an image.²⁰ To evaluate the internal architecture of adrenal lesions, TA can be used to quantify intratumoral heterogeneity based on the distribution of gray-level values and the spatial arrangement of the pixels within a given region of interest.²⁰ Previous

studies have demonstrated that TA helps distinguish among adrenal lesions. Shi et al.¹ performed TA using CT to distinguish ADM from benign adrenal masses; they used support vector machine modeling and reported an AUC of 0.85 ± 0.03 and accuracy of 77% for distinguishing metastatic from benign adrenal masses. Yi et al.²¹ used a texture software program (MaZda, version 4.6) to analyze CT images to differentiate LPA from APH. Their logistic regression model based on four texture parameters differentiated LPA from APH, with a sensitivity, specificity, and accuracy of 86.2%, 97.5%, and 94.4%, respectively. Sensitivity and specificity were higher in the present study; the difference may be related to the differences in the software used and the higher number of second-order parameters in our study. Elmohr et al.²² used TA in 54 cases of adrenal adenomas and carcinomas. Using Boruta random forest modeling, they achieved a validated accuracy of 82% for differentiating benign from malignant adrenal lesions. Ho et al.²³ performed a multivariate logistic analysis of 21 texture features, which were combined for each modality, to show that contrast-enhanced CT and chemical-shift MRI could identify malignant lesions, whereas unenhanced CT could not, in 23 patients (malignant adrenal lesions in 8 and LRAs in 15). We included more patients in our study and a significant difference was found between adenomas and malignant lesions. Torresan et al.²⁴ used a cutoff mean densitometry value of 22.5 HU for contrast-enhanced CT TA with a sensitivity and specificity of 95% and 100%, respectively.

To differentiate LRAs from other adrenal lesions (ADMs and APHs) using TA, 10 first-order and 16 second-order texture parameters were identified as statistically significant in our study. HU-Skewness, GLCM-Homogeneity, and NGLDM-Contrast were the most significantly different texture parameters between LRA and ADM. LRAs are easily differentiated from other lipid-poor adrenal lesions using conventional adrenal imaging techniques.

It is often difficult to differentiate LPA from ADM or APH in daily radiological practice. LPA is a benign condition that does not require surgical treatment. In contrast to previous studies, we focused on comparisons of first-order and second-order texture parameters to distinguish among LPA, ADM, and APH.

In our study, LPA was significantly different from ADM in appearance, based on 12 texture parameters in unenhanced CT im-

ages, with high sensitivity and specificity of 85.7% and 70.3%, respectively. The first-order texture parameters GLRLM-LGE, GLRLM-GLNU, NGLDM-Busyness, GLZLM low gray-level zone emphasis, and GLZLM-GLNU calculated on enhanced CT images had positive correlations with ADM. Moreover, HU_{mean} (OR: 3.87), GLRLM-LGE (OR: 2.64), and NGLDM-Busyness (OR: 5.07) were the texture parameters that best differentiated LPA from ADM. Our study included a sufficient number of cases to evaluate TA using 45 parameters and to construct a logistic model based on these texture parameters, allowing us to predict malignancy using TA.

We also evaluated the relationship between LPA and APH. A comparison of our results with those in recent studies that evaluated CT for the same clinical questions revealed higher sensitivity and specificity values for our TA results than those identified by Yi et al.²¹ In this study, sensitivity and specificity for distinction between LPA and APH were 86.2% and 97.5%, respectively. In addition, our model predicts APH with high accuracy using conventional assessment methods in unenhanced CT. Moreover, the GLRLM-GLNU value in the second-order TA was a statistically significant factor differentiating LPA from APH, with a high OR of 119.09.

Finally, we also evaluated the role of TA in differentiating ADM from APH in unenhanced CT images. Our logistic regression analysis identified an HU peak sphere of 0.5 mL (OR: 4.2) as a more robust factor differentiating these two lesion types compared with the other texture parameters. Our study has several limitations. First, the retrospective study design was potentially subject to bias. Previous studies have indicated that the reproducibility and variability of TA are affected by the CT device and technique used. To overcome these limitations, we included 140 patients who underwent unenhanced CT using the same acquisition protocol performed by two different CT scanners. Additional studies involving a larger number of cases and standardized CT acquisition protocols will provide more reliable results. Second, we did not include any primary malignant adrenal tumors. Although malignant adrenal tumors are rare, ADM being the most common, the inclusion of malignant adrenal tumors of different histopathological types, especially adrenocortical cancer, is needed to improve the reliability of our results. Another limitation is that we used two-dimensional analysis, which may have resulted in less available information during the data transfer compared with three-dimensional analysis.²⁵ Fi-

nally, our machine learning program used logistic regression analysis and SPSS software, whereas recent studies have used additional texture parameters and specialized machine learning programs (i.e., random forest, GBM), which can improve the results.

In conclusion, TA parameters in unenhanced CT images can be used to differentiate among the four most frequently encountered adrenal tumors, LRA, LPA, ADM, and APH. Differences in texture parameters may be related to differences in tumor cellularity, lipid content, and biological behavior. The addition of TA parameters to CT findings may improve the ability of radiologists to distinguish LPA from ADM or APH. Unenhanced CT-based TA has potential for the diagnosis of incidental adrenal lesions, leading to accurate diagnoses and preventing unnecessary surgical treatment. Further prospective trials in larger populations are needed to verify the role and performance of TA in patients with adrenal masses.

Conflict of interest disclosure

The authors declared no conflicts of interest.

References

1. Shi B, Zhang GM, Xu M, Jin ZY, Sun H. Distinguishing metastases from benign adrenal masses: what can CT texture analysis do? *Acta Radiol.* 2019;60(11):1553-1561. [\[CrossRef\]](#)
2. Song JH, Chaudhry FS, Mayo-Smith WW. The incidental indeterminate adrenal mass on CT (> 10 H) in patients without cancer: is further imaging necessary? Follow-up of 321 consecutive indeterminate adrenal masses. *AJR Am J Roentgenol.* 2007;189(5):1119-1123. [\[CrossRef\]](#)
3. Schieda N, Siegelman ES. Update on CT and MRI of Adrenal Nodules. *AJR Am J Roentgenol.* 2017;208(6):1206-1217. [\[CrossRef\]](#)
4. Albano D, Agnello F, Midiri F, et al. Imaging features of adrenal masses. *Insights Imaging.* 2019;10(1):1. [\[CrossRef\]](#)
5. Bharwani N, Rockall AG, Sahdev A, et al. Adrenocortical carcinoma: the range of appearances on CT and MRI. *AJR Am J Roentgenol.* 2011;196(6):W706-W714. [\[CrossRef\]](#)
6. Leung K, Stamm M, Raja A, Low G. Pheochromocytoma: the range of appearances on ultrasound, CT, MRI, and functional imaging. *AJR Am J Roentgenol.* 2013;200(2):370-378. Erratum in: *AJR Am J Roentgenol.* 2013;200(3):705. [\[CrossRef\]](#)
7. Johnson PT, Horton KM, Fishman EK. Adrenal imaging with multidetector CT: evidence-based protocol optimization and interpretative practice. *Radiographics.* 2009;29(5):1319-1331. Erratum in: *Radiographics.* 2009;29(7):2198. [\[CrossRef\]](#)
8. Schieda N, Al Dandan O, Kielar AZ, Flood TA, McInnes MD, Siegelman ES. Pitfalls of adrenal imaging with chemical shift MRI. *Clin Radiol.* 2014;69(11):1186-1197. [\[CrossRef\]](#)
9. Choi YA, Kim CK, Park BK, Kim B. Evaluation of adrenal metastases from renal cell carcinoma and hepatocellular carcinoma: use of delayed contrast-enhanced CT. *Radiology.* 2013;266(2):514-520. [\[CrossRef\]](#)
10. Kato H, Kanematsu M, Zhang X, et al. Computer-aided diagnosis of hepatic fibrosis: preliminary evaluation of MRI texture analysis using the finite difference method and an artificial neural network. *AJR Am J Roentgenol.* 2007;189(1):117-122. [\[CrossRef\]](#)
11. Basara Akin I, Ozgul H, Simsek K, Altay C, Secil M, Balci P. Texture analysis of ultrasound images to differentiate simple fibroadenomas from complex fibroadenomas and benign phyllodes tumors. *J Ultrasound Med.* 2020;39(10):1993-2003. [\[CrossRef\]](#)
12. Koçak B, Durmaz EŞ, Ateş E, Kılıçkesmez Ö. Radiomics with artificial intelligence: a practical guide for beginners. *Diagn Interv Radiol.* 2019;25(6):485-495. [\[CrossRef\]](#)
13. Xu B, Gao J, Cui L, et al. Characterization of adrenal metastatic cancer using FDG PET/CT. *Neoplasma.* 2012;59(1):92-99. [\[CrossRef\]](#)
14. Mazzaglia PJ, Monchik JM. Limited value of adrenal biopsy in the evaluation of adrenal neoplasm. *Arch Surg.* 2009;144(5):465-470. [\[CrossRef\]](#)
15. Nioche C, Orlhac F, Boughdad S, et al. LIFEX: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res.* 2018;78(16):4786-4789. [\[CrossRef\]](#)
16. Johnson PT, Horton KM, Fishman EK. Adrenal mass imaging with multidetector CT: pathologic conditions, pearls, and pitfalls. *Radiographics.* 2009;29(5):1333-1351. [\[CrossRef\]](#)
17. Patel J, Davenport MS, Cohan RH, Caoili EM. Can established CT attenuation and washout criteria for adrenal adenoma accurately exclude pheochromocytoma? *AJR Am J Roentgenol.* 2013;201(1):122-127. [\[CrossRef\]](#)
18. Seo JM, Park BK, Park SY, Kim CK. Characterization of lipid-poor adrenal adenoma: chemical-shift MRI and washout CT. *AJR Am J Roentgenol.* 2014;202(5):1043-1050. Erratum in: *AJR Am J Roentgenol.* 2019;212(1):232. [\[CrossRef\]](#)
19. Elbanan MG, Javadi S, Ganeshan D, et al. Adrenal cortical adenoma: current update, imaging features, atypical findings, and mimics. *Abdom Radiol (NY).* 2020;45(4):905-916. [\[CrossRef\]](#)
20. Ursprung S, Beer L, Bruining A, et al. Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma—a systematic review and meta-analysis. *Eur Radiol.* 2020;30(6):3558-3566. [\[CrossRef\]](#)
21. Yi X, Guan X, Chen C, et al. Adrenal incidentaloma: machine learning-based quantitative texture analysis of unenhanced CT can effectively differentiate sPHEO from lipid-poor adrenal adenoma. *J Cancer.* 2018;9(19):3577-3582. [\[CrossRef\]](#)
22. Elmohr MM, Fuentes D, Habra MA, et al. Machine learning-based texture analysis for differentiation of large adrenal cortical tumours on CT. *Clin Radiol.* 2019;74(10):818.e1-818.e7. [\[CrossRef\]](#)
23. Ho LM, Samei E, Mazurowski MA, et al. Can texture analysis be used to distinguish benign from malignant adrenal nodules on unenhanced CT, contrast-enhanced CT, or in-phase and opposed-phase MRI? *AJR Am J Roentgenol.* 2019;212(3):554-561. [\[CrossRef\]](#)
24. Torresan F, Crimi F, Geccato F, et al. Radiomics: a new tool to differentiate adrenocortical adenoma from carcinoma. *BJS Open.* 2021;5(1):zraa061. [\[CrossRef\]](#)
25. Soni N, Priya S, Bathla G. Texture analysis in cerebral gliomas: a review of the literature. *AJNR Am J Neuroradiol.* 2019;40(6):928-934. [\[CrossRef\]](#)

Supplementary Table 1. Abdominal CT examination protocols used in patients included in the study

	CT scanner I	CT scanner II
CT scanner brand	Brilliance 16; Philips Medical Systems, best the Netherlands	Brilliance 64; Philips Medical Systems, best the Netherlands
Reconstructed slice thickness in 3 plane	2.0 mm	2.0 mm
Tube voltage	120	120
mA	Automatic tube current modulation activated, 100–350 mA	Automatic tube current modulation activated, 100–350 mA
Detector configuration	16 x 1.25	64 x 0.625
Detector collimation	4 x 1.25	16 x 0.625
Effective section thickness	1.25 mm	1.24 mm
Reconstruction interval	2 mm	2 mm
Pitch	1.375	1.75
CT, computed tomography.		