Published in final edited form as:

Acad Radiol. 2014 December ; 21(12): 1587–1596. doi:10.1016/j.acra.2014.07.023.

CT Texture Analysis of Renal Masses:

Pilot Study Using Random Forest Classification for Prediction of Pathology

Siva P. Raman, MD, Yifei Chen, BS, James L. Schroeder, PhD, Peng Huang, PhD, and Elliot K. Fishman, MD

Department of Radiology, JHOC 3251, Johns Hopkins University, 601 N. Caroline Street, Baltimore, MD 21287 (S.P.R., Y.C., J.L.S., E.K.F.) and Biostatistics and Bioinformatics Division, Department of Oncology, Johns Hopkins University, Baltimore, Maryland (P.H.)

Abstract

Rationale and Objectives—Computed tomography texture analysis (CTTA) allows quantification of heterogeneity within a region of interest. This study investigates the possibility of distinguishing between several common renal masses using CTTA-derived parameters by developing and validating a predictive model.

Materials and Methods—CTTA software was used to analyze 20 clear cell renal cell carcinomas (RCCs), 20 papillary RCCs, 20 oncocytomas, and 20 renal cysts. Regions of interest were drawn around each mass on multiple slices in the arterial, venous, and delayed phases on renal mass protocol CT scans. Unfiltered images and spatial band-pass filtered images were analyzed to quantify heterogeneity. Random forest method was used to construct a predictive model to classify lesions using quantitative parameters. The model was externally validated on a separate set of 19 unknown cases.

Results—The random forest model correctly categorized oncocytomas in 89% of cases (sensitivity = 89%, specificity = 99%), clear cell RCCs in 91% of cases (sensitivity = 91%, specificity = 97%), cysts in 100% of cases (sensitivity = 100%, specificity = 100%), and papillary RCCs in 100% of cases (sensitivity = 100%, specificity = 98%).

Conclusions—CTTA, in conjunction with random forest modeling, demonstrates promise as a tool to characterize lesions. Various renal masses were accurately classified using quantitative information derived from routine scans.

Keywords

Texture analysis; clear cell renal cell carcinoma; papillary renal cell carcinoma; oncocytoma; multidetector computed tomography

© AUR, 2014

Address correspondence to: S.P.R. sraman3@gmail.com.

Disclosures and conflicts of interest: The authors have no relevant financial disclosures or relevant conflicts of interest.

Disclosures: CTTA was performed on commercially available software (TexRAD Ltd, Somerset, UK). The software vendor provided training to all the authors before the study. None of the authors have any financial interest in TexRAD Ltd.

When confronted with an indeterminate renal mass on computed tomography (CT), the ability of a radiologist to accurately differentiate common histologic subtypes (such as clear cell renal cell carcinoma [RCC], papillary RCC, chromophobe RCC, oncocytomas, etc.) is limited to either an analysis of morphologic features or Hounsfield attenuations. Many studies have used some combinations of enhancement characteristics and lesion morphology to differentiate these masses, and although these various tumor types as *groups* tend to have different appearances and enhancement patterns, there is a significant overlap between lesion categories that prevents prospective prediction of a lesion's underlying histology with high confidence (1–10). As a result, solid renal masses are usually surgically resected unless the patient is a poor surgical candidate or the lesion measures less than 2 cm in size. If more confident prospective characterization were possible, it is conceivable that lesions that were strongly thought to represent more indolent variants of RCC (ie, papillary and chromophobe RCC) or benign oncocytomas could be followed rather than resected.

Computed tomography texture analysis (CTTA) is a quantitative technique that allows users to characterize heterogeneity within a region of interest (ROI) based on the distribution of pixel intensities and gray-level values using both unfiltered and frequency filtered images by deriving quantitative texture parameters based on attributes of the pixel values themselves and the image histogram. This quantitative technique has been primarily used in a number of studies as a means of predicting patient outcomes and prognosis (11–26). However, there has been only limited application of this method toward lesion characterization and the differentiation of lesions with similar radiographic appearance (such as various types of solid renal masses), and although the few works that have dealt with this topic have demonstrated quantitative differences in texture variables between different lesions, they have not sought to create true *predictive models* using texture data (27–30).

This preliminary work on CTTA seeks to apply texture analysis to the differentiation and classification of a few common types of renal masses, including oncocytomas, clear cell RCC, papillary RCC, and renal cysts. The goal of this pilot study is to assess the efficacy of texture analysis, when combined with a robust statistical classification model, in differentiating this small group of renal masses, and thereby evaluate the promise of texture analysis as a quantitative imaging tool that might ultimately be applied to a larger number of lesion categories (31).

MATERIALS AND METHODS

Patient Selection

Approval was obtained from the Institutional Review Board for this retrospective study, and patient informed consent was waived. Health Insurance Portability and Accountability Act (HIPAA) compliance was maintained throughout the study.

An internal Pathology Department database was searched for surgically resected oncocytomas, papillary RCCs, and clear cell RCCs. Consecutive cases were selected from patients in the database with the following inclusion criteria: (1) patients had preoperative imaging performed *at our institution* between 2008 and 2013; (2) multiphase imaging was performed using a dedicated renal mass protocol; and (3) lesions measured at least 2 cm in

size in all three dimensions (to obtain meaningful information from each of the texture analysis software's spatial filters). The exclusion criterion was adequate representation of the lesion type had already been achieved. Definition of adequacy is detailed in the "Power analysis" section.

A further search was conducted on a Radiology Department database of renal protocol CT scans performed at our institution during the same period (2008–2014) for patients with benign-appearing renal simple cysts. The inclusion criteria for this subset of patients were as follows: (1) patients had a simple renal cyst measuring at least 2 cm in all three dimensions and (2) the cyst demonstrated no internal complexity, calcification, or enhancement in any phase of the scan (Bosniak I classification). Exclusion criteria for this subset of patients were as follows: (1) cysts were adjacent (<1 cm) to an enhancing mass; (2) the cyst had indeterminate boundaries; and (3) adequate representation of the patient subset had been achieved. Notably, the presence of another renal tumor in the same kidney as the cyst was not an exclusion criterion, and some of the patients with oncocytoma or RCC in the pathology cohort mentioned previously had simple cysts that met inclusion criteria.

To achieve equivalent weighting of each of the previously described lesion classes (oncocytoma, papillary RCC, clear cell RCC, and simple cyst) in the subsequent statistical analysis, an equal number of lesions for each class were selected ($n = 20$ lesions of each type) to be used in the creation of the predictive model. If multiple discrete lesions were present in a single patient these lesions were counted individually, although treated as "clustered" in the subsequent statistical analysis. As a result, although there were an equal number of lesions of each class, there were different numbers of total patients used to generate the model. The sizes of the patient cohorts were as follows: 20 patients with oncocytomas, 18 patients with papillary RCCs, 20 patients with clear cell RCCs, and 12 patients with simple cysts.

A further subset of patients that were not included in the generation of the random forest model was used for external validation of the classification model. This "validation" cohort comprised 19 patients, with four oncocytomas, five papillary RCCs, five clear cell RCCs, and five simple cysts. These 19 patients were acquired from the same Pathology and Radiology Department databases as the original set of 80 patients used to construct the random forest model, with identical inclusion and exclusion criteria. Baseline characteristics of each of the patient cohorts are shown in Table 1.

CT Technique

All CT examinations were acquired on one of two multidetector CT scanners (Siemens Medical Solutions, Malvern, PA, USA) in use at our institution from 2008 to 2013: (1) Siemens Somatom Sensation 64 (detector collimation 64×0.6 mm, reconstruction at 3-mm slice thickness and 3-mm slice intervals, 120 kVp, 150–200 mAs), or (2) Siemens Somatom Definition Flash dual-source (detector collimation 128×0.6 mm, reconstruction at slice thickness of 3 mm and 3-mm slice intervals, 120 kVp, quality reference 290 mAs for online dose modulation system [CareDose 4D; Siemens Medical Solutions]).

All cases were acquired using a standardized renal mass protocol that did not change during the time course of the study, as follows: after the acquisition of noncontrast images, arterial, venous, and excretory phase images were acquired at 25–30, 60–70, and 240 seconds, respectively, after the intravenous injection of contrast. Subsequently, the arterial, venous, and excretory phase images were all used for CTTA. Notably, noncontrast images were not used for analysis, as lesion boundaries could not be accurately discerned for many cases, and the CTTA software package did not allow an ROI to be propagated from one phase to another (coregistration was not possible). Contrast used was either iohexol (Omnipaque 350; GE Healthcare) or iodixanol (Visipaque 320; GE Healthcare, Pewaukee, WI, USA) infused through a peripheral intravenous line at 3–5 mL/s and water as a oral contrast agent. The attenuation value of the aorta at the level of the renal arteries was measured for each phase of each study, and was subsequently used in the statistical analysis.

CT Texture Analysis

All analysis was performed using a commercially available CTTA software product (Version 1.1; TexRAD Ltd, Somerset, UK), for which dedicated training was provided to the study members by the manufacturer before the study, and which was purchased by our institution. Archived arterial, venous, and delayed phase images were anonymized and uploaded to a remote server, allowing analysis by the software package. Using the software package, an ROI was manually drawn around the margins of the renal mass, and the software automatically applied threshold to exclude pixels with Hounsfield attenuation values less than –50, thereby excluding gas or fat at the margins of the mass. Polygonal ROIs were drawn in conjunction by two observers, the more senior who had 3 years of experience as a body imaging attended, whereas the more junior observer was a postbaccalaureate research assistant provided with specific training in evaluating the kidneys on CT. The research assistant drew all the ROIs once the boundaries of lesions had been decided by the more senior observer, and the accuracy of every ROI was confirmed by the more senior observer before incorporation of each case into the model. In addition, the senior observer also made the final decision regarding the simple cyst cases to be included in the study. ROIs were selected from an equal number of slices in each of the three phases being used for CTTA (arterial, venous, and delayed). Hand-selected polygonal ROIs were drawn around the lesions in each of the three phases using roughly equal boundaries and matched slices as shown in Figure 1.

To capture a representative group of texture values for each patient, ROIs were obtained from a maximum of 10 axial slices for each mass. For lesions that were large in the craniocaudal dimension (ie, >10 slices in the craniocaudal dimension), 10 consecutive slices were selected from the tumor. Notably, every slice was not analyzed in these large masses to avoid *overweighting* any single lesion in the subsequent statistical model. Similarly, a minimum of five slices was required for any given mass to avoid *underweighting* any given mass in the model. Importantly, the used software package only allowed analysis of axial slices (and consequently only two-dimensional filtration), and as a result, no volumetric analysis could be performed. For each study, the ROI was drawn on the lesion using roughly equal boundaries and from an equal number of anatomically matched slices in each phase.

The clear cell RCCs had an average maximum axial diameter of 43.0 mm (range, 20–74 mm), with an average of 1933 pixels in each arterial phase ROI, 2046 pixels in each venous phase ROI, and 1914 pixels in each delayed phase ROI. The mean attenuation values for the clear cell RCCs were 82.4 HU on the arterial phase, 82.7 HU on the venous phase, and 63.91 HU on the delayed phase images. The papillary RCCs had an average maximum axial diameter of 57.3 mm (range, 12–154 mm), with an average of 5606 pixels in each arterial phase ROI, 5798 pixels in each venous phase ROI, and 5304 pixels in each delayed phase ROI. The mean attenuation values for the papillary RCCs were 38.5 HU on the arterial phase, 47.0 HU on the venous phase, and 43.8 HU on the delayed phase images. The oncocytomas had an average maximum axial diameter of 41.6 mm (range, 20–107 mm), with an average of 1890 pixels in each arterial phase ROI, 1901 pixels in each venous phase ROI, and 1791 pixels in each delayed phase ROI. The mean attenuation values for the oncocytomas were 75.8 HU on the arterial phase, 99.3 HU on the venous phase, and 68.6 HU on the delayed phase images. The renal cysts had an average maximum axial diameter of 44 mm (range, 25–74 mm), with an average of 1650 pixels in each arterial phase ROI, an average of 1781 pixels in each venous phase ROI, and an average of 1669 pixels in each delayed phase ROI. The mean attenuation values for the cysts were 6.8 HU on the arterial phase, 9.0 HU on the venous phase, and 8.4 HU on the delayed phase images.

Although the basic methodology of CTTA has been described in great detail elsewhere, the process can be briefly described as follows (13–21,30,32): once ROIs were acquired, the CTTA software filtered each slice using several Laplacian of Gaussian spatial band-pass filters at a range of spatial periods (ie, 1/spatial frequency). Five filters, preset by the software developer and termed as spatial scale filters of 2, 3, 4, 5, and 6 mm, were applied to each slice. Each of these filters created a derived image, which highlighted the features of the slice at the corresponding spatial period. In addition to the filtered images, unfiltered images (ie, spatial scale filters = 0) from each slice were also used for analysis. As has been previously described elsewhere, once the images were filtered, the software package calculated a set of predetermined quantitative parameters including mean gray-level pixel intensity (average brightness), entropy (irregularity), standard deviation of pixel intensity (width of the distribution of intensities), kurtosis (peakedness of distribution), and skewness (asymmetry of distribution) from the selected ROIs (13–21,30,32). These five variables were acquired from each filter (along with the unfiltered data), for a total of 30 CTTA variables (=five parameters $\times$ six images analyzed) and normalized to the contrast load by using the measured enhancement of the aorta from each phase. Additionally, age and gender were also used in the statistical model.

Statistics

The data were exported from the CTTA software and analyzed using both commercial and open source statistics software packages (S-Plus; TIBCO Software, Palo Alto, CA and R, <http://www.r-project.org>).

A “random forest” method was used to build a predictive model to classify lesion types (31,33,34). This “machine-learning” method uses a combination of multiple “classification and regression trees” that are independent prediction algorithms with decision points

selected on an automated and randomized basis. Briefly, the CTTA data, along with the known classification of each mass based on pathologic diagnosis (for resected lesions) or radiographic appearance (for simple cysts) was fed to the algorithm along with each patient's age and sex. Each decision tree was constructed using a bootstrapped sample (ie, a resampled subset of data which generally included two-third of the data set), with the unselected data (usually one-third of the data) set aside for subsequent error prediction. These samples that were not included in the "bootstrap sample" were termed the "out-of-bag" (OOB) data, and were later used to internally validate the accuracy of the derived random forest model.

The bootstrap sample was used to create a classification tree, with different variables used at each decision branch point. The bootstrap sampling consists of two stages to handle correlated image slices from the same lesion. The first stage was lesion-based sampling, and the second stage was slice-based sampling (from selected lesions). Different variables were used at each decision branch. The variable to be used at each branch point was selected at random, and the cutoff for that variable was selected to optimize classification of the bootstrapped data. The limits for each tree were increased up to the maximum size of the variable data structure, and no "pruning" optimization was performed. This process of bootstrapping data and generating decision trees was repeated 5000 times, and the resulting group of decision trees is referred to as the random forest.

After generation of the random forest, the model can be used for prospective classification of data corresponding to unknown lesion types by running the data from each prospective case through each of the decision trees individually and taking the mode output (ie, majority vote among all trees) as the model's prediction for the corresponding prospective case. This prospective performance of the random forest model can be internally estimated using lesions not selected by the bootstrap sampling, even if prospective or external cases are not available, using a technique called OOB error calculation. This error prediction is the standard way to assess the performance of a random forest model. In theory, the use of OOB data to estimate an error rate should produce results similar to the error rate calculated from an independent validation data set. For the OOB calculation, a classification prediction is made for each slice by each decision tree using only data from lesions that were excluded from the bootstrap data set used to generate each respective tree. Because each tree is thus only using data for which its decision points were never tested during tree generation, any potential "overfitting" error as a result of testing previously modeled data is avoided. The performance of the random forest model is displayed as "confusion" matrices showing the rate of correct classification based on the OOB test data. Mode class was also used to predict a lesion using classifications of all slices from the same lesion.

Another available output of the random forest model is quantitative comparison of the different variables used in the model on the basis of their importance to classification performance (ie, "variable importance" plots). The method underlying this comparison is described elsewhere (31); it is briefly based on the percent increase in misclassification rate when that variable's values are permuted randomly compared to the misclassification rate with that variable's values left intact.

Power Analysis

Notably, a conventional power analysis was not performed before the study. The CTTA process generates a large number of quantitative image-derived parameters, but there was no hypothesis as to the expected difference of each of these parameters between the various lesion types. Furthermore, power calculations are not typically performed with random forest modeling; the predicted performance of the model is the only criteria used to establish the quality of the model's fit. The considerations used to determine the size of the study cohort were to include a balanced group size of each potential lesion type as to achieve equal weighting within the model for each group. The number of cases (20) was chosen based on the limited number of available cases of oncocytoma within the pathology database (24), some of which were randomly selected to remain as prospective validation cases not used to create the initial models.

RESULTS

Random Forest Model

Error rates (based on the OOB data) for the random forest prediction models are summarized in Table 2. Table 2 is a confusion matrix detailing the accuracy of four separate random forest models on a “per-slice” basis: a model using data from all three contrast phases (arterial, venous, and delayed) in conjunction, and then three separate models using only a single contrast phase (arterial, venous, or delayed). As expected, the best performance is seen with a model using data from all three contrast phases in conjunction. Interestingly, of the models using only a single contrast phase in isolation, the model derived from the arterial phase data demonstrates the best performance, whereas the venous and delayed phases demonstrate considerably higher error rates (with the delayed phase providing the most poor results). Notably, regardless of the phases of contrast used, the biggest sources of error are present in the model's ability to discriminate between clear cell RCC and oncocytoma.

Variable Importance Plots

Variable importance plots for each of the aforementioned four random forest models are illustrated in Figures 2–5. These plots represent statistical prioritization of the many variables used to construct each of the random forest models in terms of their contribution to the predictive model. As described previously, the “ranking” of each variable within these plots is directly proportional to the magnitude of improvement conferred by that variable to the accuracy of the model. In each of the four models, mean and standard deviation, both in the filtered and unfiltered data, were the most important variables.

Performance of Model Using Validation Data Set

Table 3 is a confusion matrix demonstrating the accuracy (on a per-slice basis) of the four random forest models for the 19 cases (113 slices), which were not used in the generation of the random forest model, and which served as the independent validation data set for the external validation of the model. Most importantly, when all three contrast phases were used in conjunction, the model correctly classified 108 of 113 slices (error rate of 4.4%).

Moreover, as the theory underlying random forest modeling would predict, the error rates calculated from the OOB data set were not appreciably different from the error rates based on the validation data set.

If the classification of each case was allowed to be the *mode* class (ie, majority vote of the classifications for each slice within a lesion) of all the slices analyzed, the model incorporating all three phases in conjunction *would correctly classify all 19 lesions*. However, the model using arterial phase data alone would misclassify a single oncocytoma as a clear cell RCC, correctly categorizing 18 of 19 cases, the model using venous phase data alone would misclassify a single oncocytoma as a clear cell RCC, four clear cell RCCs as oncocytomas, and a single cyst as a papillary RCC, correctly categorizing 13 of 19 cases, and the model using delayed phase data alone would misclassify a single papillary RCC as an oncocytoma, a single oncocytoma as a clear cell RCC, a single clear cell RCC as an oncocytoma, and a single clear cell RCC as a papillary RCC, correctly categorizing 15 of 19 cases.

DISCUSSION

Traditional evaluation of a renal mass using CT and prediction of the underlying histology are largely based on the radiologist's visual examination of tumor morphology and grossly quantifying lesion enhancement through Hounsfield attenuation values. There are certain broad patterns that can suggest one tumor versus another: clear cell RCCs are usually avidly enhancing lesions with high tumor-to-cortex enhancement ratios, and these masses often demonstrate substantial washout (often >40%) between the arterial and delayed phases. Alternatively, papillary RCCs tend to be more hypovascular with lower tumor-to-cortex and tumor-to-aorta ratios, as well as a lack of significant washout between the arterial and delayed phases. Finally, although oncocytomas also tend to be relatively vascular masses that can often resemble clear cell RCCs, studies have suggested that these lesions may have slightly lesser degrees of enhancement compared to clear cell RCCs on both corticomedullary and excretory phase images (7,10). Certain morphologic features may also be helpful, including the presence of a central scar, calcification, and intratumoral cystic change. Nevertheless, there is a significant overlap between different lesions with regard to both enhancement and morphology, and solid renal masses without intratumoral fat are almost always surgically resected because of the inability to distinguish benign masses (such as oncocytoma) from those that are malignant (ie, RCC) and aggressive subtypes of RCC (ie, clear cell RCC) from more indolent lesions (ie, papillary or chromophobe RCC) (1–7,35–40).

However, texture analysis, which describes lesion heterogeneity through an analysis of the distribution of pixel intensities and positioning of pixels within an ROI at different spatial frequencies, might allow the quantification of lesion attributes, which may not be readily apparent to the radiologist either based on visual inspection or Hounsfield attenuation analysis, and could theoretically be helpful in classification of lesion types. However, until recently, only a small number of studies had evaluated the utility of CTTA for this purpose, including a study by Ba-Ssalamah et al. (29) evaluating various gastric masses, a study by Huang et al. (27) seeking to differentiate benign hemangiomas from various malignant

hepatic masses, and a study by Kido et al. (28) evaluating various types of lung lesions. However, these studies were all limited to evaluating pairwise comparisons of various texture parameters, without a true attempt to create either a predictive model or use more complex statistical techniques taking into different *combinations* of texture parameters to differentiate lesion types.

As our previous work has demonstrated, simply looking at *individual* texture parameters in isolation is not useful when seeking to classify lesions (30). Any given parameter, in isolation, is not useful for lesion discrimination, and moreover, there is often not a statistically significant difference between different lesion types with regard to an individual parameter. However, *random forest analysis* is very well suited to the data extracted from CTTA, as demonstrated in this pilot study. *Random forest*, a technique first described by Breiman, is a machine-learning technique that allows a sample to be classified into a few known classes using a *hierarchy* of variables. Each “forest” is composed of a large number of randomly generated “trees,” each of which categorizes the data set using a random subset of different variables, and a majority “vote” of all trees is used to classify any given sample (31,41).

Table 2 illustrates the accuracy of the model in differentiating the four lesion types when using data from all three phases (arterial, venous, and delayed). Each classification tree in the random forest is created based on a bootstrap sample comprising roughly two-third of the observations in the data set, with the remaining one-third comprising the OOB data. These OOB data are used to generate the error estimates for the statistical model, which prior studies have shown to be roughly equivalent to true prospective validation. In a model using all three contrast phases in conjunction, 100% of papillary RCCs and renal cysts were correctly categorized with error rates of only 8.75% and 11% for clear cell RCCs and oncocytomas, respectively. In other words, oncocytomas and clear cell RCCs could be correctly differentiated in roughly 90% of cases. We subsequently tested our model on a separate set of 19 cases that were *not* used in the generation of the random forest model. Of the 113 analyzed slices from these 19 lesions, 108 were correctly categorized, and if the classification result was deemed to be the mode class of all the slices analyzed for each lesion, all 19 cases were correctly categorized into the four lesion types. As other studies detailing random forest technique have suggested, our accuracy rates based on the OOB data and the prospective data set are roughly equivalent. Table 3 also details the performance of models using only the arterial, venous, or delayed phases in isolation, which all show inferior results to a model using all three phases in conjunction. Interestingly, of these three models, the model derived from the arterial phase data alone was the most accurate.

The data from this preliminary study suggest that CTTA might offer promise as a quantitative imaging tool that may someday augment a radiologist’s ability to characterize lesion histology. The ability to better differentiate renal masses prospectively without the need for biopsy could potentially impact patient management: lesions strongly thought to represent either a benign entity (such as an oncocytoma) or a more indolent form of RCC (such as a papillary RCC) could theoretically be followed sequentially over time. Moreover, as percutaneous biopsy of renal masses has become more common over the last few years, CTTA could serve as a valuable adjunct, particularly in light of problems with non-

diagnostic biopsies and sampling error that are relatively common with percutaneous biopsy. Notably, the ability to accurately distinguish clear cell RCCs and oncocytomas is of vital importance before surgery, and our current inability to confidently distinguish these two entities results in the routine resection of most hypervascular renal masses. The random forest model was able to distinguish these two entities accurately in roughly 90% of cases, and it is conceivable that the model might be even more accurate if a larger number of cases were used to generate the model. In addition, the high accuracy (100%) of the model in differentiating papillary RCCs from simple renal cysts is also worth noting. Although these two entities are not typically difficult to differentiate if Hounsfield attenuations are carefully measured, papillary RCCs are not uncommonly incorrectly categorized as cysts when Hounsfield attenuations values are not carefully measured, particularly when only one phase of contrast is available for review. CTTA would seem to have promise in eliminating this mistake, as despite the two lesions' superficial similarities, cysts and papillary RCCs produce completely different texture parameters.

Nevertheless, despite the promising results of this pilot study, a few limitations and future study directions should certainly be acknowledged: (1) this is merely a pilot study examining a small subset of renal lesions to evaluate the efficacy of this quantitative method when combined with random forest statistical analysis. At this time, there is certainly not sufficient validation of this technique to suggest active clinical use of this method in day-to-day practice, and such a clinical application would require robust studies incorporating many more different types of renal masses, far larger sample sizes, and more rigorous prospective validation. This study is merely meant to illustrate the early promise of this method for this indication and suggests the need for further rigorous scientific evaluation. (2) The CTTA software we used provided preset filters (with specific spatial periods) which could not be altered. As a result, to acquire meaningful results at spatial periods up to 6 mm, lesions were included only if they measured 2 cm in size in all three dimensions. Clearly, as many lesions less than 2 cm are increasingly being followed over time rather than resected, CTTA would be increasingly valuable if its accuracy could be validated for smaller lesions. (3) Only four common renal masses were studied, including Bosniak I cysts, clear cell RCCs, papillary RCCs, and oncocytomas. There are certainly other relatively common lesions which would warrant inclusion of a larger validation study, including lipid poor angiomyolipomas, hyperdense renal cysts, and chromophobe RCCs. Moreover, it is not uncommon for pathologic analysis of these lesions to reveal overlapping histology (ie, RCC with chromophobe and papillary components). This study excluded such cases from analysis, a potential limitation in real-life practice. (4) The software package we used applied threshold to eliminate pixels with Hounsfield attenuation values less than -50 to eliminate the unintentional inclusion of fat or gas at the margins of the ROI, particularly given that ROI selection was manual. It is possible that this might affect the texture parameters for any given lesion with internal fat (which can be present with both adenomas and hepatocellular carcinoma (HCC)). (5) Although our sample size (80 cases for generation of the random forest model and 19 cases for prospective validation) is certainly as large (or larger) than other comparable studies, any use of CTTA in routine clinical practice would require validation using much larger numbers of analyzed cases, particularly given the diversity of imaging appearances the different studied renal lesions can have. (6) Finally, it should be

noted that both the texture variables acquired and the statistical modeling technique used are nonintuitive in the sense that a direct correlation between the visually apparent CT features and the acquired texture variables/statistical modeling results cannot be made. As a result, it is difficult to draw strong conclusions about why the model performs better with some lesions rather than others.

CONCLUSIONS

CTTA allows the quantification of lesion heterogeneity based on the distribution of pixel intensities within an ROI. In this preliminary study, when combined with random forest statistical modeling, this technique allowed relatively accurate discrimination and characterization of a small subset of common renal masses, suggesting its future promise as a quantitative imaging tool that may augment our ability to accurately predict a lesion's histology based on its imaging appearance. In this study, based on CTTA parameters acquired from arterial, venous, and delayed phase images, a random forest model was able to correctly categorize clear cell RCCs in 92.5% of cases, papillary RCCs in 100% of cases, oncocytomas in 89% of cases, and cysts in 100% of cases. Moreover, when the model was tested on an additional set of 19 cases not included in the original random forest model, the model correctly categorized all 19 cases. The results of this pilot study suggest the promise of this technique and argue for further studies with larger numbers of lesion types, larger sample sizes, and more rigorous prospective validation.

Acknowledgments

Statistical analysis collaboration was supported by grants #P50CA103175 and #P30CA006973 from the National Institutes of Health.

References

1. Sheir KZ, El-Azab M, Mosbah A, et al. Differentiation of renal cell carcinoma subtypes by multislice computerized tomography. *J Urol*. 2005; 174:451–455. [PubMed: 16006863]
2. Kim JH, Bae JH, Lee KW, et al. Predicting the histology of small renal masses using preoperative dynamic contrast-enhanced magnetic resonance imaging. *Urology*. 2012; 80:872–876. [PubMed: 22854134]
3. Jung SC, Cho JY, Kim SH. Subtype differentiation of small renal cell carcinomas on three-phase MDCT: usefulness of the measurement of degree and heterogeneity of enhancement. *Acta Radiol*. 2012; 53:112–118. [PubMed: 22114020]
4. Zhang J, Lefkowitz R, Ishill N, et al. Solid renal cortical tumors: differentiation with CT. *Radiology*. 2007; 244:494–504. [PubMed: 17641370]
5. Choi SK, Jeon SH, Chang SG. Characterization of small renal masses less than 4 cm with quadruphasic multidetector helical computed tomography: differentiation of benign and malignant lesions. *Korean J Urol*. 2012; 53:159–164. [PubMed: 22468210]
6. Jinzaki M, Tanimoto A, Mukai M, et al. Double-phase helical CT of small renal parenchymal neoplasms: correlation with pathologic findings and tumor angiogenesis. *J Comput Assist Tomogr*. 2000; 24:835–842. [PubMed: 11105696]
7. Young JR, Margolis D, Sauk S, et al. Clear cell renal cell carcinoma: discrimination from other renal cell carcinoma subtypes and oncocytoma at multiphasic multidetector CT. *Radiology*. 2013; 267:444–453. [PubMed: 23382290]

8. Herts B, Coll D, Novick A, et al. Enhancement characteristics of papillary renal neoplasms revealed on triphasic helical CT of the kidneys. *AJR Am J Roentgenol.* 2002; 178:367–372. [PubMed: 11804895]
9. Bird VG, Kanagarajah P, Morillo G, et al. Differentiation of oncocytoma and renal cell carcinoma in small renal masses (<4 cm): the role of 4-phase computerized tomography. *World J Urol.* 2011; 29:787–792. [PubMed: 20717829]
10. Raman SP, Johnson PT, Allaf ME, et al. Chromophobe renal cell carcinoma: multiphase MDCT enhancement patterns and morphologic features. *AJR Am J Roentgenol.* 2013; 201:1268–1276. [PubMed: 24261366]
11. Al-Kadi OS, Watson D. Texture analysis of aggressive and nonaggressive lung tumor CE CT images. *IEEE Trans Biomed Eng.* 2008; 55:1822–1830. [PubMed: 18595800]
12. Davnall F, Yip CS, Ljungqvist G, et al. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights Imaging.* 2012; 3:573–589. [PubMed: 23093486]
13. Ganeshan B, Abaleke S, Young RC, et al. Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage. *Cancer Imaging.* 2010; 10:137–143. [PubMed: 20605762]
14. Ganeshan B, Miles KA, Young RC, et al. In search of biologic correlates for liver texture on portal-phase CT. *Acad Radiol.* 2007; 14:1058–1068. [PubMed: 17707313]
15. Ganeshan B, Miles KA, Young RC, et al. Hepatic enhancement in colo-rectal cancer: texture analysis correlates with hepatic hemodynamics and patient survival. *Acad Radiol.* 2007; 14:1520–1530. [PubMed: 18035281]
16. Ganeshan B, Miles KA, Young RC, et al. Texture analysis in non-contrast enhanced CT: impact of malignancy on texture in apparently disease-free areas of the liver. *Eur J Radiol.* 2009; 70:101–110. [PubMed: 18242909]
17. Ganeshan B, Miles KA, Young RC, et al. Three-dimensional textural analysis of brain images reveals distributed grey-matter abnormalities in schizophrenia. *Eur Radiol.* 2010; 20:941–948. [PubMed: 19760235]
18. Ganeshan B, Panayiotou E, Burnand K, et al. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *Eur Radiol.* 2012; 22:796–802. [PubMed: 22086561]
19. Ganeshan B, Skogen K, Pressney I, et al. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol.* 2012; 67:157–164. [PubMed: 21943720]
20. Ganeshan B, Strukowska O, Skogen K, et al. 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. [PubMed: 22649761]
21. Goh V, Ganeshan B, Nathan P, et al. Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker. *Radiology.* 2011; 261:165–171. [PubMed: 21813743]
22. Miles KA, Ganeshan B, Griffiths MR, et al. Colorectal cancer: texture analysis of portal phase hepatic CT images as a potential marker of survival. *Radiology.* 2009; 250:444–452. [PubMed: 19164695]
23. Ng F, Ganeshan B, Kozarski R, et al. 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. [PubMed: 23151829]
24. Radulescu E, Ganeshan B, Minati L, et al. Gray matter textural heterogeneity as a potential in-vivo biomarker of fine structural abnormalities in As-perger syndrome. *Pharmacogenomics J.* 2013; 13:70–79. [PubMed: 22333911]
25. Ravanelli M, Farina D, Morassi M, et al. Texture analysis of advanced non-small cell lung cancer (NSCLC) on contrast-enhanced computed tomography: prediction of the response to the first-line chemotherapy. *Eur Radiol.* 2013; 23:3450–3455. [PubMed: 23835926]
26. Skogen K, Ganeshan B, Good C, et al. Measurements of heterogeneity in gliomas on computed tomography relationship to tumour grade. *J Neuro-oncol.* 2013; 111:213–219.

27. Huang YL, Chen JH, Shen WC. Diagnosis of hepatic tumors with texture analysis in nonenhanced computed tomography images. *Acad Radiol*. 2006; 13:713–720. [PubMed: 16679273]
28. Kido S, Kuriyama K, Higashiyama M, et al. Fractal analysis of small peripheral pulmonary nodules in thin-section CT: evaluation of the lung-nodule interfaces. *J Comput Assist Tomogr*. 2002; 26:573–578. [PubMed: 12218822]
29. Ba-Ssalamah A, Muin D, Scherthaner R, et al. Texture-based classification of different gastric tumors at contrast-enhanced CT. *Eur J Radiol*. 2013; 82:e537–e543. [PubMed: 23910996]
30. Raman S, Schroeder J, Huang P, et al. Classification of hypervascular liver lesions using CT texture analysis: generation of a predictive model. *J Comput Assist Tomogr*. in press.
31. Breiman L. Random forests. *Machine Learning*. 2001; 45:5–32.
32. Ganeshan B, Burnand K, Young R, et al. Dynamic contrast-enhanced texture analysis of the liver: initial assessment in colorectal cancer. *Invest Radiol*. 2011; 46:160–168. [PubMed: 21102348]
33. Ho TK. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell*. 1998; 20:832–844.
34. Kleinberg E. An overtraining-resistant stochastic modeling method for pattern recognition. *Ann Statist*. 1996; 24:2319–2349.
35. Kang SK, Chandarana H. Contemporary imaging of the renal mass. *Urol Clin North Am*. 2012; 39:161–170. vi. [PubMed: 22487759]
36. Kim J, Kim T, Ahn H, et al. Differentiation of subtypes of renal cell carcinoma on helical CT scans. *AJR Am J Roentgenol*. 2002; 178:1499–1506. [PubMed: 12034628]
37. Sauk SC, Hsu MS, Margolis DJ, et al. Clear cell renal cell carcinoma: multiphasic multidetector CT imaging features help predict genetic karyotypes. *Radiology*. 2011; 261:854–862. [PubMed: 22025734]
38. Shebel H, Elsayes K, Sheir KZ, et al. Quantitative enhancement washout analysis of solid cortical renal masses using multidetector computed tomography. *J Comput Assist Tomogr*. 2011; 35:337–342. [PubMed: 21586926]
39. Takagi T, Kondo T, Tanabe K. Impact of the tumor enhancement pattern in computed tomography for the differential diagnosis of renal cell carcinoma and benign renal tumor. *Int J Urol*. 2011; 18:866–867. [PubMed: 21992007]
40. Zhang C, Li X, Hao H, et al. The correlation between size of renal cell carcinoma and its histopathological characteristics: a single center study of 1867 renal cell carcinoma cases. *BJU Int*. 2012; 110:E481–E485. [PubMed: 22540142]
41. Amit Y, Geman D. Shape quantization and recognition with randomized trees. *Neural Comput*. 1997; 9:1545–1588.

	Arterial	Venous	Delay
Oncocytoma
Clear Cell RCC
Papillary RCC
Simple Cyst

Figure 1.
Examples of the region of interest selection and matching for different lesions on the arterial, venous, and delayed phase images.

Figure 2.

Variable importance plot for three phases (arterial, venous, delayed) in conjunction of the constructed random forest model. The plot details the relative importance of each of the quantitative texture parameters (which is listed on the y-axis of the plot) to the model's accuracy. The x-axis is the mean decrease in the Gini co-efficient that results when that variable is included in the model. The Gini coefficient is a measure of inequality among the trees in the random forest, and in this case represents the performance of the random forest model with and without a variable included. Those variables that have the highest decrease in the Gini coefficient were most likely to create consensus among the individual decision trees used in the model (or reduce inequality) when included in the model. These variables are therefore most predictive of the outcome of the model overall. Those variables with a small decrease in the mean Gini coefficient are relatively less important to the prediction made by the random forest model.

Figure 3.

Variable importance plot for the arterial phase of the random forest model. The plot details the relative importance of each of the quantitative texture parameters (which is listed on the y-axis of the plot) to the model's accuracy. The x-axis is the mean decrease in the Gini coefficient that results when that variable is included in the model.

Figure 4.

Variable importance plot for the venous phase of the random forest model. The plot details the relative importance of each of the quantitative texture parameters (which is listed on the y-axis of the plot) to the model's accuracy. The x-axis is the mean decrease in the Gini coefficient that results when that variable is included in the model.

Figure 5.

Variable importance plot for the delayed phase of the random forest model. The plot details the relative importance of each of the quantitative texture parameters (which is listed on the y-axis of the plot) to the model's accuracy. The x-axis is the mean decrease in the Gini coefficient that results when that variable is included in the model.

TABLE 1

Demographic and Patient Information for the 70 Patients (80 Lesions) Used in the Construction of the Random Forest Models and the 18 Patients (19 Lesions) Used for External Validation of the Models

	Oncocytoma	Clear Cell	Papillary	Simple Cyst
Random Forest Modeled Subset				
Number of Patients	20	20	18	12
Average Age (y)	66	65	64	65
Female (%)	25%	50%	28%	42%
Total Number of Lesions	20	20	20	20
Total Number of Slices	154	160	158	150
Average Number of Slices/Lesion	7.7	8.0	7.9	7.5
Validation Subset				
Number of Patients	4	4	5	5
Average Age (y)	66	47	65	60
Female (%)	25%	25%	0%	40%
Total Number of Lesions	4	5	5	5
Total Number of Slices	21	31	36	25
Average Number of Slices/Lesion	5.3	6.2	7.2	5.0

TABLE 2

Confusion Matrices for Random Forest Models (Based on the Out-of-Bag Data) Using all Three Contrast Phases in Conjunction and Models Using Only the Arterial, Venous, or Delayed Phases Alone

Lesion Type	Oncocytoma	Clear Cell RCC	Papillary RCC	Simple Cyst	Sensitivity (%)	Specificity (%)	Total Class Error (%)
3-Phase							
Oncocytoma	137	16	1	0	89	99	11.0
Clear Cell RCC	7	146	7	0	91	97	8.8
Papillary RCC	0	0	158	0	100	98	0.0
Simple Cyst	0	0	0	150	100	100	0.0
Arterial							
Oncocytoma	136	15	3	0	88	98	11.7
Clear Cell RCC	10	145	5	0	91	97	9.4
Papillary RCC	1	0	156	1	99	98	1.3
Simple Cyst	0	0	1	149	99	100	0.7
Venous							
Oncocytoma	142	12	0	0	92	97	7.8
Clear Cell RCC	16	136	7	1	85	97	15.0
Papillary RCC	0	3	153	2	97	98	3.2
Simple Cyst	0	0	1	149	99	99	0.7
Delayed							
Oncocytoma	125	23	5	1	81	94	18.8
Clear Cell RCC	24	132	4	0	83	94	17.5
Papillary RCC	2	4	151	1	96	98	4.4
Simple Cyst	0	0	1	149	99	100	0.7

RCC, renal cell carcinoma.

The results are presented on a per-slice basis. The “true” pathologic diagnosis is represented by the heading for each row, whereas the classification provided by the model is represented by the heading for each column.

TABLE 3
Confusion Matrices Detailing the Accuracy of the Random Forest Models when Analyzing the 19 Validation Cases

	Oncocytoma	Clear Cell RCC	Papillary RCC	Simple Cyst	Sensitivity (%)	Specificity (%)	Total Class Error (%)
3-Phase							
Oncocytoma	18	3	0	0	85.7	97.8	14.3
Clear Cell RCC	2	29	0	0	93.5	96.3	6.5
Papillary RCC	0	0	36	0	100	100	0.0
Simple Cyst	0	0	0	25	100	100	0.0
Arterial							
Oncocytoma	17	4	0	0	81.0	97.8	19.0
Clear Cell RCC	2	29	0	0	93.5	95.1	6.5
Papillary RCC	0	0	36	0	100	100	0.0
Simple Cyst	0	0	0	25	100	100	0.0
Venous							
Oncocytoma	19	2	0	0	90.5	73.9	9.5
Clear Cell RCC	24	7	0	0	22.6	97.6	77.4
Papillary RCC	0	0	36	0	100	97.4	0.0
Simple Cyst	0	0	2	23	92.0	100	8.0
Delayed							
Oncocytoma	18	3	0	0	85.7	83.7	14.3
Clear Cell RCC	11	15	5	0	48.4	93.9	51.6
Papillary RCC	4	2	30	0	83.3	93.5	16.7
Simple Cyst	0	0	0	25	100	100	0.0

RCC, renal cell carcinoma.

The results are presented on a per-slice basis. The true pathologic diagnosis is represented by the heading for each row, whereas the classification provided by the model is represented by the heading for each column.