



Lexicon for renal mass terms at CT and MRI: a consensus of the society of abdominal radiology disease-focused panel on renal cell carcinoma

Atul B. Shinagare^{1,2,3} · Matthew S. Davenport^{3,4} · Hyesun Park^{1,2} · Ivan Pedrosa^{3,5} · Erick M. Remer^{3,6} · Hersh Chandarana^{3,7} · Ankur M. Doshi^{3,7} · Nicola Schieda^{3,8} · Andrew D. Smith^{3,9} · Raghunandan Vikram^{3,10} · Zhen J. Wang^{3,11} · Stuart G. Silverman^{1,3}

Received: 30 April 2020 / Revised: 25 June 2020 / Accepted: 4 July 2020 / Published online: 18 August 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Purpose There is substantial variation in the radiologic terms used to characterize renal masses, leading to ambiguity and inconsistency in clinical radiology reports and research studies. The purpose of this study was to develop a standardized lexicon to describe renal masses at CT and MRI.

Materials and methods This multi-institutional, prospective, quality improvement project was exempt from IRB oversight. Thirteen radiologists belonging to the Society of Abdominal Radiology (SAR) disease-focused panel on renal cell carcinoma representing nine academic institutions participated in a modified Delphi process to create a lexicon of terms used to describe imaging features of renal masses at CT and MRI. In the first round, members voted on terms to be included and proposed definitions; subsequent voting rounds and a teleconference established consensus. One non-voting member developed the questionnaire and consolidated responses. Consensus was defined as $\geq 80\%$ agreement.

Results Of 37 proposed terms, 6 had consensus to be excluded. Consensus for inclusion was reached for 30 of 31 terms (13/14 basic imaging terms, 8/8 CT terms, 6/6 MRI terms and 3/3 miscellaneous terms). Despite substantial initial disagreement about definitions of ‘renal mass,’ ‘necrosis,’ ‘fat,’ and ‘restricted diffusion’ in the first round, consensus for all was eventually reached. Disagreement remained for the definition of ‘solid mass.’

Conclusions A modified Delphi method produced a lexicon of preferred terms and definitions to be used in the description of renal masses at CT and MRI. This lexicon should improve clarity and consistency of radiology reports and research related to renal masses.

Keywords Renal mass · Renal cell carcinoma · CT, MRI, lexicon, Delphi method

✉ Atul B. Shinagare
ashinagare@bwh.harvard.edu

¹ Department of Radiology, Brigham and Women’s Hospital/
Harvard Medical School, 75 Francis Street, Boston,
MA 02215, USA

² Dana-Farber Cancer Institute, Boston, USA

³ Society of Abdominal Radiology Disease-Focused Panel on
Renal Cell Carcinoma, Houston, TX, USA

⁴ Departments of Radiology & Urology, Michigan Medicine,
Ann Arbor, MI, USA

⁵ University of Texas Southwestern, Dallas, TX, USA

⁶ Cleveland Clinic, Cleveland, OH, USA

⁷ NYU Langone Health, New York, NY, USA

⁸ University of Ottawa, Ottawa, ON, Canada

⁹ University of Alabama at Birmingham, Birmingham, AL,
USA

¹⁰ The University of Texas M.D. Anderson Cancer Center,
Houston, TX, USA

¹¹ UCSF, San Francisco, CA, USA

Introduction

Diagnostic radiologists translate imaging data into words that influence patient care [1]. These words must be precise and accurate to optimize communication between radiologists, referring providers, and patients [2–10]. Emerging data indicate the need for structured reports [11–15] that use specific well-defined terms [16–21] to improve clarity and understandability. To address these goals, schemas such as the Breast Imaging Reporting and Data System (BI-RADS) [2, 3] have created standardized lexicons that define terms to be used when radiologists are describing imaging findings, either in a clinical radiology report or in a research study. Such lexicons define terms and provide formal recommendations about which specific terms should be used or avoided [12, 13].

No such lexicon exists for CT and MRI of renal masses. Terminology used in radiology reports and research studies remains highly variable [22–25]. For example, numerous terms have been used to describe the presence of ‘fat’ in a renal mass, including terms that refer to fat cells (e.g., ‘macroscopic fat,’ ‘macroscopic lipid,’ ‘extracellular fat,’ ‘bulk fat’) and terms that refer to both fat cells and other cells that contain fat (e.g., ‘microscopic fat,’ ‘intracellular fat,’ ‘intracellular lipid,’ ‘intracytoplasmic fat,’ ‘intra-voxel fat’) [26–29]. Terms used inconsistently or inaccurately, or that are not based on explicit definitions, lead to confusing or ambiguous clinical reports [29], interobserver variability, and difficulty interpreting and reproducing research studies [23–25].

There is a need to develop standard terms for imaging features of renal masses that will promote accurate and consistent reporting in both clinical practice and research. The purpose of this study was to develop a standardized lexicon and image atlas to describe renal masses at CT and MRI.

Methods

This study was undertaken by the Society of Abdominal Radiology (SAR) Disease-Focused Panel (DFP) on renal cell carcinoma (RCC). Only terms relevant to CT and MRI were included because CT and MRI are the most common modalities used to evaluate indeterminate renal masses. Terms describing imaging protocols (e.g., the phases of enhancement after intravenous contrast material administration), anatomic structures (e.g., renal cortex), and specific disease entities (e.g., renal cell carcinoma) were not evaluated because they are outside the scope of this project and definitions for these terms and concepts exist throughout the literature. In addition, some terms that are

specific to cystic masses (e.g., ‘nodule,’ ‘septa’ and the adjectives used to describe them) were excluded from this project because they were being fully defined as part of a parallel effort to update the Bosniak classification [30]. All 13 radiologist members of the RCC DFP from nine academic institutions in the US and Canada participated in the study. All participants considered renal masses a primary area of clinical and research interest.

Study design: modified Delphi method

This multi-institutional, prospective, quality improvement project was exempt from IRB oversight, and utilized a modified Delphi method [31–34]. One radiologist member of the panel (AS), with the help of a radiology clinical fellow (HP) served as the ‘coordinator’ who compiled the initial list of terms, prepared the questionnaires, and collected and analyzed the data. The remaining twelve radiologist members of the RCC DFP were invited to participate in the creation of the lexicon. All members are fellowship-trained abdominal radiologists with mean 13 years of experience as attending radiologist (range

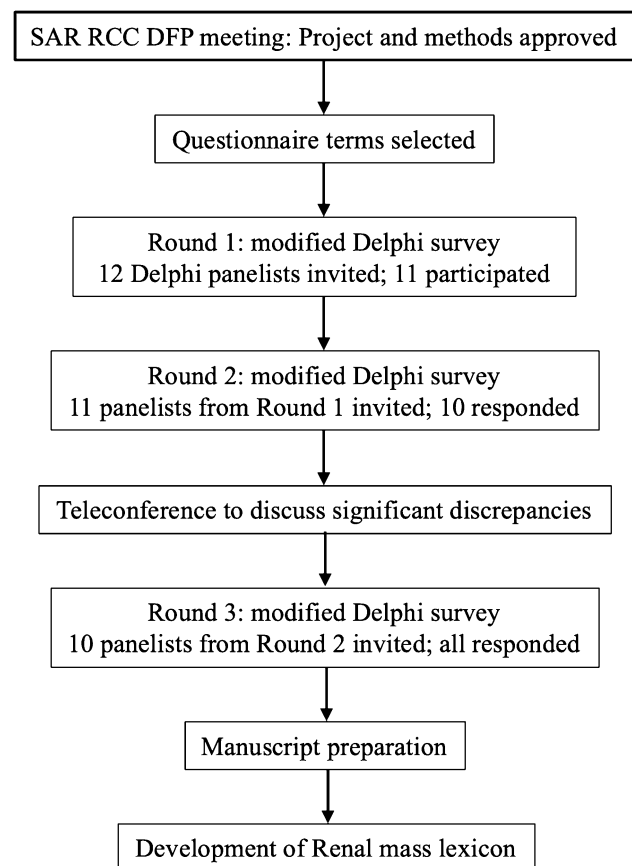


Fig. 1 Flowchart demonstrating the modified Delphi process used to create a lexicon of terms for the description of renal masses at CT and MRI

5–30 years). Figure 1 outlines the process that was followed, and Appendix 1 provides the detailed methodology. Three rounds of questionnaires and one teleconference (after the second round) were conducted to build consensus (Appendix 1). Following the third round of questionnaires, the manuscript was created. Blinded edits to the manuscript were made and final consensus was reached. Based on prior literature, $\geq 80\%$ agreement at the end of three rounds was considered sufficient ‘consensus’ regarding the inclusion or exclusion of each term and its definition [35–38]. Individual panelist responses remained anonymous in all three rounds. After the lexicon was finalized, official endorsement was obtained from the SAR Board of Directors.

Results

The Round 1 questionnaire was sent to 12 panelists by one coordinator; 11 completed it. Of the 11 panelists who participated in Round 1, one was not able to participate further; as a result, ten panelists completed the questionnaires in Rounds 2 and 3.

A total of 35 terms were included in the first round (Table 1). There were 17 basic imaging terms, eight CT terms, six MRI terms, and four miscellaneous terms. Two additional terms were proposed in the first round (‘magnetic susceptibility’ and ‘growth rate’). By the end of the third round, of the 37 terms, 31 had $\geq 80\%$ agreement to

remain in the lexicon (Table 2) and six had $\geq 80\%$ agreement to be removed from the lexicon (Table 3). Of the 31 included terms, 29 had 100% agreement to include and two had 90% agreement to include. Of the six excluded terms, four had 100% agreement to exclude, one had 90% agreement to exclude, and one had 80% agreement to exclude (Table 3).

Of 31 included terms, consensus on definition was achieved for 30 (13/14 basic imaging terms, 8/8 CT terms, 6/6 MRI terms, and 3/3 miscellaneous terms). Of these, 17 terms (10 basic imaging terms, four CT terms, and three MRI terms) had 100% agreement, 11 terms (two basic imaging terms, four CT terms, two MRI terms, and three miscellaneous terms) had 90% agreement, and two terms (one basic imaging term and one MRI term) had 80% agreement. Consensus was not reached for the definition of one general term (solid mass, 60% agreement). Table 2 details the proposed definitions and points of disagreement. The lexicon was endorsed by the SAR Board of Directors.

Discussion

We developed a consensus-based lexicon to describe renal masses at CT and MRI that addresses a knowledge gap in clinical radiology reporting. As radiology practice becomes more value-based, there is an increasing emphasis on the quality and actionability of radiology reports. This renal

Table 1 Initial list of imaging terms evaluated during the first round of a modified Delphi process to create a lexicon of standard terms to be used in the description of renal masses at CT and MRI

Basic imaging terms (<i>n</i> = 17)	CT terms (<i>n</i> = 8)	MRI terms (<i>n</i> = 6):	Miscellaneous terms (<i>n</i> = 4):
Renal mass	Attenuation	Signal intensity	Heterogeneity
Composition	Internal reference for attenuation	Fat at MRI	Rate of enhancement
Solid mass	Fluid attenuation	Calcification at MRI	Degree of enhancement
Cystic mass	Macroscopic fat at CT	Hemorrhage at MRI	Internal vascularity
Size	Calcification at CT	Enhancement at MRI	
Laterality	Morphologic categories of calcification	Restricted (impeded) diffusion	
Margin	Hemorrhage at CT		
Well-defined margin	Enhancement at CT		
Ill-defined margin			
Outline			
Growth pattern			
Categories of growth pattern			
Anteroposterior location			
Craniocaudal location			
Necrosis			
Capsule			
Central scar			

Table 2 Final consensus ($\geq 80\%$ agreement) terms and definitions to be used in the description of renal masses at CT and MRI

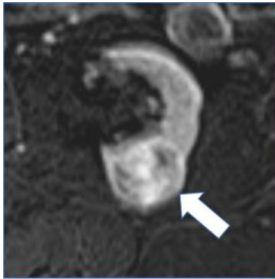
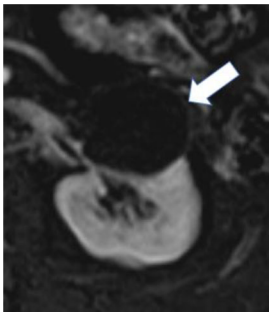
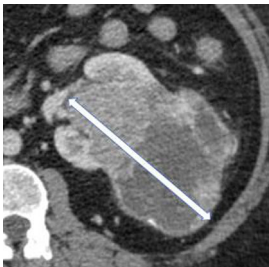
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
Basic imaging terms	Renal mass	100	100	Any space-occupying abnormality thought to originate from or contiguous with renal parenchyma. Abnormalities include space-occupying features which deviate from the normal appearance of renal parenchyma and surrounding tissues	One panelist agreed with the proposed definition but also recommended that benign entities with pathognomonic appearance, such as hypertrophic column of Bertin, should not be considered a renal mass
	Composition	100	100	The internal make-up of a renal mass, which may be further categorized as solid or cystic	
	Solid mass	100	60	Mass composed of greater than or equal to 25% solid enhancing components or fat	Four (of 10) panelists commented that solid component does not always enhance and may appear as heterogeneous non-enhancing tissue
					
	Cystic mass	100	90	Mass composed of less than 25% enhancing components or fat	
					
	Size	100	100	The single longest dimension of a renal mass in any plane	
					
Laterality	100	100	Localization of renal mass(es) in the right, left or bilateral kidneys		
Margin	100	100	The interface between the mass and adjacent uninvolved tissue		

Table 2 (continued)

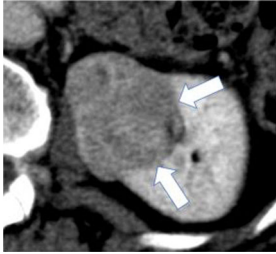

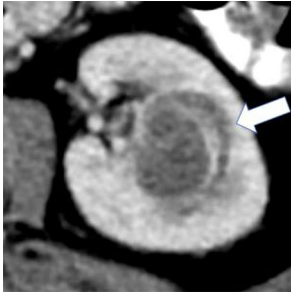

Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	Well-defined margin 	100	90	A sharp pencil-thin demarcation between the mass and surrounding tissue over at least 90% circumference of the mass	One panelist suggested using sharp demarcation over “most of the circumference of the mass” instead of using 90% threshold
	Ill-defined margin 	100	100	An indistinct or blurred demarcation between the mass and surrounding tissue over greater than 10% circumference of the mass	
	Growth pattern	100	100	The location of a renal mass relative to the renal parenchymal contour	
	Categories of growth pattern -Endophytic  - < 50% Exophytic  - ≥ 50% Exophytic	100	100	Growth pattern may be categorized into: 1. Endophytic: Located within the renal parenchyma such that the outer renal contour is not altered. 2. < 50% Exophytic: The renal contour is altered and < 50% of the mass projects out from the renal contour. 3. ≥ 50% Exophytic: The renal contour is altered and ≥ 50% of the mass projects out from the renal contour	

Table 2 (continued)

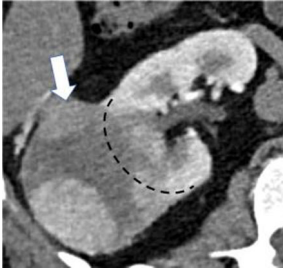
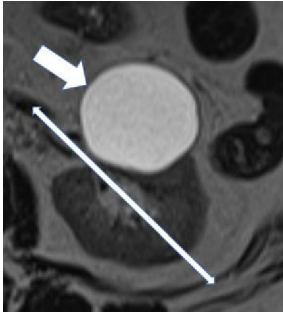
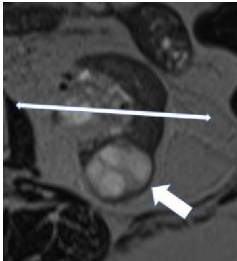
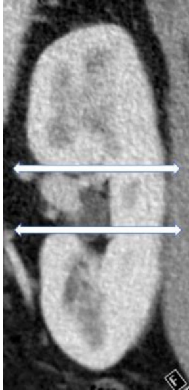
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	 <p>Anteroposterior location - Anterior</p>	100	100	The anterior or posterior location of a mass within the renal parenchyma on axial images. The mass is 'anterior' if the center of the mass lies anterior to the line bisecting the renal hilum on axial images and 'posterior' if the center of the mass lies posterior to this line	
	 <p>- Posterior</p>				
	 <p>Craniocaudal location</p>	100	100	The location of a mass with respect to the polar lines. A renal mass can be centered in the upper pole, interpolar region or lower pole	
					

Table 2 (continued)


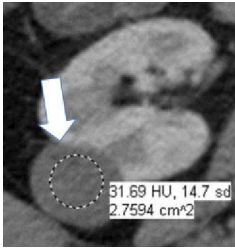
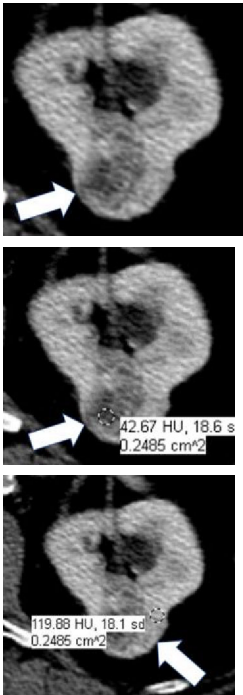
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	<p>Necrosis</p> 	100	80	<p>Necrosis is a pathology term and cannot be diagnosed with certainty at imaging; necrosis, fibrosis, and cystic change can appear similar. Necrosis may be considered when non-enhancing ill-defined components of variable CT attenuation or MRI signal intensity are identified</p>	
CT terms	<p>Attenuation</p> <p>Homogenous mass</p>  <p>Heterogenous mass</p> 	100	100	<p>CT attenuation can be determined in one or more regions of a renal mass. If the mass is homogeneous, the attenuation can be obtained with a single region of interest (ROI) that includes at least 2/3rds of the mass, being careful not to include tissues outside the mass. If the mass is heterogeneous, multiple smaller ROIs should be placed, taking care not to volume-average areas of different attenuation within the mass. Since attenuation may change depending on the phase of enhancement, reference to CT attenuation should be clear about when attenuation is being measured</p>	

Table 2 (continued)

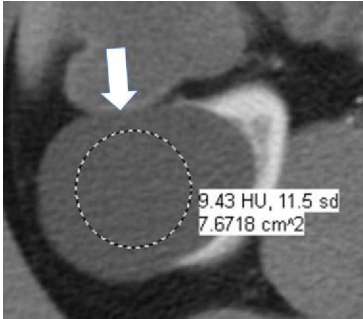
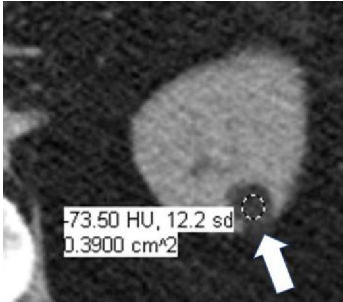
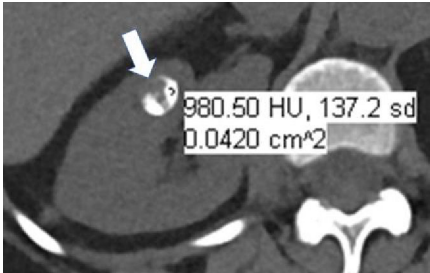
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	Internal reference for attenuation	100	100	Renal mass attenuation (e.g., classified as hypo-, iso- or hyperattenuating) is generally considered relative to renal cortex. However, reference standard anatomy should be specified in the report. When using the term 'attenuation' without referencing renal cortex or other reference standard, 'hypoattenuating', 'isoattenuating', and 'hyperattenuating' refer to a comparison to simple fluid (-9 to 20 HU)	
	Fluid attenuation	100	100	Non-enhancing content within a mass measuring -9 HU to + 20 HU	
					
	Macroscopic fat at CT	100	100	Attenuation -10 HU or less (not reaching attenuation of air which is close to -1000). Pixel-based region of interest assessments will measure quantum mottle and are not reliable. Presence of fat cannot be confidently excluded on contrast-enhanced CT	
					
	Calcification at CT	100	90	Attenuation greater than 100 HU on unenhanced CT (excluding metal). Smaller calcifications may have lower attenuation because of volume averaging. Calcification cannot be confidently excluded on contrast-enhanced CT	
					

Table 2 (continued)


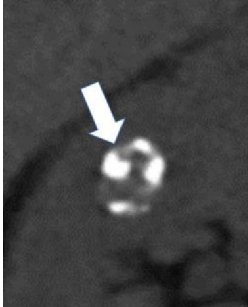
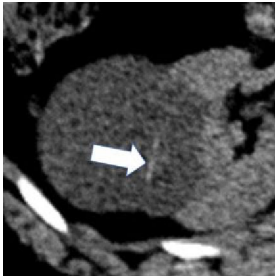
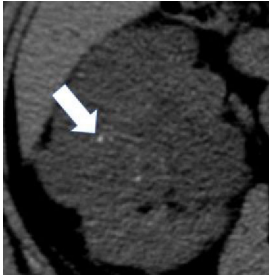
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	Morphologic categories of calcification - Thin	100	90	Calcification may be further characterized as thin or thick, linear or punctate, border-forming or non-border-forming	Clinical significance of morphologic appearance of calcification within a renal mass is controversial
					
	- Thick				
					
	- Linear				
					
	- Punctate				
					
	- Border-forming				

Table 2 (continued)



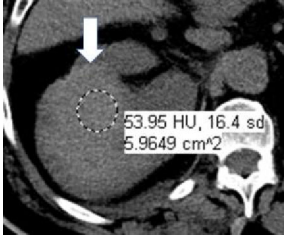
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	 <p data-bbox="288 617 456 642">- Non-border-forming</p>  <p data-bbox="288 926 435 951">Hemorrhage at CT</p> 	100	90	<p>Hyperattenuating nonenhancing area(s) within or around a renal mass. Since the CT attenuation of hemorrhage is variable and dependent on factors such as hemoglobin concentration and age of hemorrhage, there is no specific attenuation range that is diagnostic of hemorrhage; however, hemorrhage typically ranges from 40–100 HU</p>	<p>One panelist suggested using attenuation range of 40–70 HU for hemorrhage at CT</p>

Table 2 (continued)

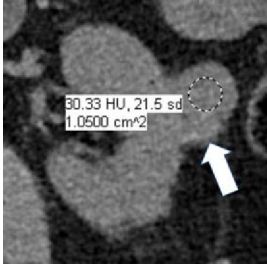
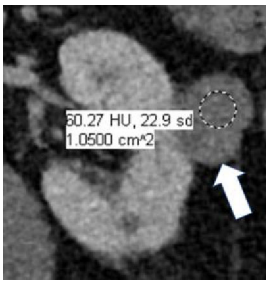
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	Enhancement at CT	100	90	<p>Increase in attenuation ≥ 20HU using region(s) of interest on any contrast-enhanced phase relative to unenhanced imaging. Equivocal enhancement is an increase in attenuation ≥ 10 HU but < 20 HU, and no enhancement is an increase in attenuation < 10 HU. The post-contrast phase should be acquired ≤ 2 min after contrast administration using the same parameters as the unenhanced imaging. When determining if a mass is enhancing at CT, the attenuation of the most enhancing portion of the mass is measured and compared to the same region on unenhanced imaging</p>	One panelist suggested removing the 2-min time limit for assessment of enhancement
		100	100	<p>MRI signal intensity ('brightness') can be determined in one or more regions of a renal mass. Since signal intensity may change depending on the sequence type and phase of enhancement, reference to signal intensity should be clear about what is being measured. The intensity is generally compared to normal renal cortex. If another reference standard is used it should be specified (such as "hyperintense compared to the muscle")</p>	

Table 2 (continued)

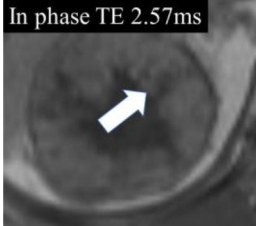

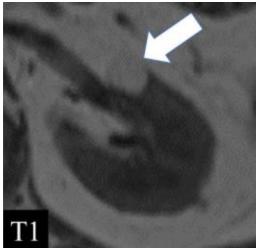
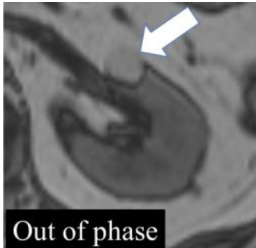
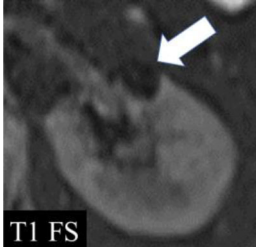
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	Fat at MRI	100	90	<p>When describing ‘fat’ at MRI, either ‘macroscopic fat’ or ‘microscopic fat’ should be used. Macroscopic fat is thought to be present when there is 1) loss of intratumoral signal intensity before and after application of fat-suppression (i.e. with no changes in any of the other acquisition parameters), or 2) linear or curvilinear chemical shift artifact causing India-ink (etching) artifact within or at the periphery of the mass at macroscopic fat–water interfaces, while the central area (i.e. inside the etching artifact) remains hyperintense (i.e., following the signal intensity of subcutaneous and intraabdominal fat). Microscopic fat is thought to be present when there is non-linear non-curvilinear focal or diffuse signal intensity loss in any portion of a renal mass on the opposed-phase relative to in-phase T1-weighted images</p>	<p>One panelist suggested that 1–19% proton density fat fraction on multi-echo DIXON acquisitions in any portion of a renal mass may indicate microscopic fat</p>
	- Microscopic				
					
					
	- Macroscopic				
					
					
					

Table 2 (continued)

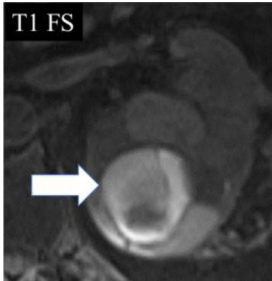

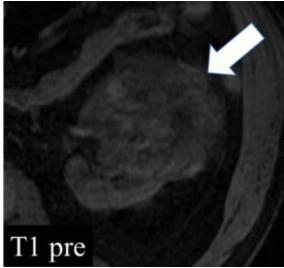
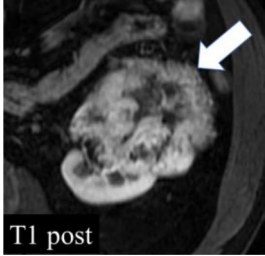
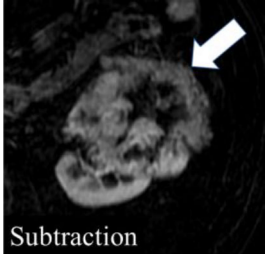
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	Hemorrhage at MRI	100	100	Non-enhancing high signal intensity (relative to renal cortex) on pre-contrast fat-suppressed T1-weighted images. Proteinaceous content can appear similar	
	 T1 FS				
	 Post contrast subtraction				
	Enhancement at MRI	100	90	Unequivocal visual enhancement or increase in signal intensity $\geq 15\%$ using region(s) of interest on any contrast-enhanced phase relative to unenhanced imaging. No enhancement is an increase in signal intensity $< 15\%$. The post-contrast phase should be determined ≤ 2 min after contrast administration using the same parameters as the unenhanced imaging. Subtraction images (post-contrast minus pre-contrast) may be used to detect enhancement qualitatively by noting any signal in a portion of a mass; however, subtraction images need to be properly registered to avoid misregistration artifact. When determining if a mass is enhancing at MRI, the signal intensity of the most enhancing portion of the mass is measured and compared to the same region on unenhanced imaging	One panelist suggested removing the 2-min time limit for assessment of enhancement
	 T1 pre				
	 T1 post				
	 Subtraction				

Table 2 (continued)

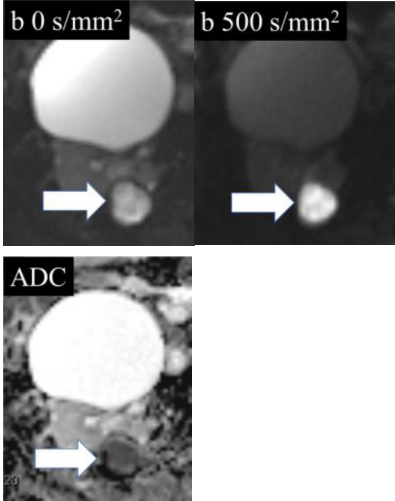
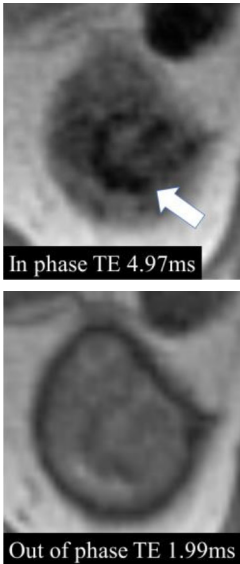
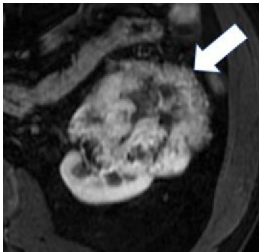

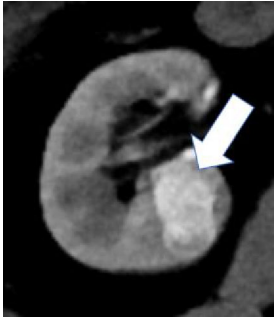
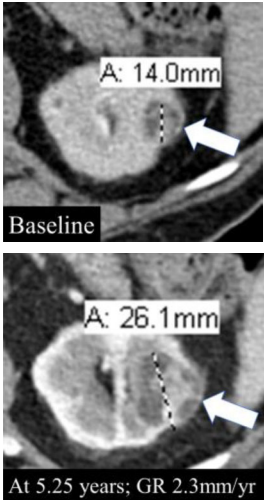
Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	Restricted (impeded) diffusion 	100	80	Relative to the renal cortex, higher signal intensity in the mass on diffusion-weighted images with b-values > 500 s/mm ² compared to a b-value of 0 s/mm ² , or lower signal intensity on apparent diffusion coefficient map relative to renal cortex	
	Magnetic susceptibility 	90	100	Magnetic susceptibility is defined as a decrease in signal intensity on MRI images acquired with a longer echo time compared to images obtained with otherwise identical acquisition parameters except for shorter echo time	One of the panelists pointed out that common causes of magnetic susceptibility artifact are hemosiderin, calcification, and gas
Miscellaneous terms	Heterogeneity 	100	90	Variable CT attenuation or MRI signal intensity. Assessment of heterogeneity may be done using visual inspection or by manually placing multiple regions of interest within a mass. There is no established quantitative definition. This definition does not include variable attenuation or signal as a result of artifacts or image noise	Heterogeneity may be assessed quantitatively using texture analysis, but there are limited data on which features to use and methods of assessment

Table 2 (continued)

Category	Term	Level of agreement for inclusion (%)	Level of agreement for definition (%)	Definition	Counterpoints and Other Comments
	Degree of enhancement	100	90		
	- Hypovascular				
					
	- Hypervascular				
					
	Growth rate	90	90	Change in size divided by the length of time over which the size changed *GR: Growth rate	
					

Arrows, solid and dotted lines on figures refer to the relevant term, each of which is defined in the column labeled ‘definition’

mass lexicon may make radiology reports of renal masses more standardized, actionable, and easy to understand for referring providers, patients, and radiologists, and may improve the reproducibility and generalizability of related research. Further, a similar lexicon may be necessary in

other clinical contexts; we also hope that this study provides a guide for similar such efforts in the future.

This lexicon addresses clinical reporting as well as the imaging-based research of renal masses. The terms and their definitions address a variety of imaging-based

Table 3 Excluded terms and level of agreement for exclusion

Category	Term	Level of agreement for the term to be excluded (%)	Reason for exclusion
Basic imaging terms	Capsule	100	Anatomic term (outside the scope of this lexicon)
Basic imaging terms	Central scar	100	Unclear clinical significance and not enough data for accurate definition
Basic imaging terms	Outline	80	Panel considered ‘margin’ to represent the same feature
MRI terms	Calcification at MR	100	Unclear clinical significance
Miscellaneous terms	Internal vascularity	100	Insufficient data for accurate definition
Miscellaneous terms	Rate of enhancement	90	Insufficient data for accurate definition

features which add clarity to radiology reports and renal mass imaging research. Some specific potential benefits include distinguishing between macroscopic and microscopic fat at MRI [26–29]. These terms reflect challenging concepts and are critical to the diagnosis of masses which contain fat cells or other cells that contain fat in their cytoplasm [39]. The lexicon also addresses imaging features used to differentiate solid and cystic masses, including the presence of intravenous contrast material enhancement [40–42]; these are important considerations when applying the Bosniak classification [30]. The lexicon does not include some of the terms used to describe cystic renal masses as the recent update proposal to the Bosniak classification provided a comprehensive description of features and definition of terms [30]. Use of the lexicon may promote consistent terminology, decrease interobserver variability, and improve reproducibility—issues that the Bosniak update proposal aimed to improve [30]. This lexicon uses consistent terminology as in the Bosniak update, but expands it to encompass all terminology related to renal masses. The lexicon also adds clarity to the description of renal mass location and growth pattern; both have implications for prognosis and management, particularly surgical planning and calculating the R.E.N.A.L nephrometry score [43].

When conducting research, investigators from different institutions do not always analyze the same imaging features. When they do, the exact terms and their definitions are either incompletely expressed or inconsistent with prior studies [22–25]. For example, in 2014 and 2015, Karlo et al. and Shinagare et al. each reported on the radiogenomics of renal cell carcinoma [23, 24]. While five of six mutations (VHL, BAP1, PBRM1, SETD2, KDM5C, MUC4) studied in these reports were the same, of the 10 imaging features evaluated, only four were evaluated in both studies. Furthermore, the conclusions reached for the

four shared imaging features also differed despite that both studies utilized similar study cohorts; this was thought to be due to differences in the definitions of imaging features used in these studies [23, 24]. For some of the imaging features studied, the definitions were unclear. For example, Shinagare et al. defined ‘necrosis’ as ‘hypodense, non-enhancing areas which were not sharply demarcated and lacked apparent walls.’ Karlo et al. defined ‘necrosis’ as “presence or absence of areas within the tumor that did not demonstrate contrast enhancement during the nephrographic and delayed phases’ [23, 24]. These studies also differed in whether specific definitions were provided for well-defined margin and tumor architecture [23, 24]. This is not meant to be a criticism of either study; indeed, during our Delphi method, there was no initial consensus for the imaging-based definition of ‘necrosis.’ Although consensus was eventually reached, the panel determined that necrosis cannot be diagnosed with certainty at imaging and therefore was difficult to define using imaging (Table 2). These difficulties highlight the need for a lexicon of standard terminology.

The modified Delphi process we used offered several advantages for arriving at expert consensus [31–34, 38]. Since each panelist completed the questionnaire independently and anonymously, their opinions were expressed freely and with sufficient time to research and formulate views, as opposed to the in-person roundtable method where a few members can dominate the discussion, and others may not feel as confident or comfortable expressing their opinions [31–34, 38]. We shared individual anonymous opinions and opposing views in Rounds 2 and 3, and geographically disparate members from different time zones discussed the sources of disagreement during the teleconference.

Given that the definitions were created using expert consensus, some of the features were by necessity

arbitrarily defined (e.g., ‘well-defined’). We were forced to rely on expert experience in the absence of published data or published definitions. There was initial disagreement on the definition of some terms, including basic terminology that is used in daily practice. ‘Renal mass’ was one of them. The panel eventually agreed that a renal mass should include any space-occupying abnormality as opposed to including only cysts and neoplasms. The panel believed that using a broad term that encompasses any pathology would decrease the chance of a non-neoplastic condition such as focal bacterial pyelonephritis from being misdiagnosed as a neoplasm. Both ‘cystic’ and ‘solid’ were also difficult to define. Although discriminating cystic masses from solid masses is a fundamental tenant in renal mass evaluation, at the time this work began, there were no established imaging criteria for doing so. This likely explains the difficulty in reaching consensus for the definition of a ‘solid’ mass. Overall, there was general agreement that a renal mass is considered solid when it is composed of $\geq 25\%$ enhancing components or fat [30]; however, there was disagreement about what constitutes solid tissue. While all the panelists agreed that a solid mass may be comprised of enhancing tissue or fat, four of 10 panelists also commented that solid tissue does not always enhance and may appear as heterogeneous non-enhancing tissue at either CT or MRI.

‘Margin’ of a renal mass is often a source of confusion. The panel agreed that a ‘well-defined margin’ denotes a sharp, pencil-thin demarcation with an abrupt transition between the mass and the surrounding tissue encompassing at least 90% of the circumference of the mass. If a well-defined margin is not present, it is termed an ‘ill-defined margin.’ There also may be confusion about the internal reference standard to assess ‘CT attenuation,’ ‘MRI signal intensity,’ and ‘enhancement.’ There was consensus to use normal renal cortex as an internal reference standard. When a renal mass is heterogeneous, one or more regions of interest may need to be placed for accurate assessment. In prior literature, several terms are variably and interchangeably used to describe the presence of fat within a renal mass [26–29]. The panel recommends use of ‘macroscopic fat’ and ‘microscopic fat,’ and provides definitions for each [39].

The lexicon does not dictate when each term should be used. The terms in the lexicon complement the structured reporting template previously derived by the SAR RCC DFP [44–46]. These studies addressed what is preferred to be included in structured reports, such as mass type (cystic mass according to Bosniak classification vs. a solid mass), presence of fat, presence of enhancement, and radiologic stage—all considered ‘core’ features when reporting indeterminate renal masses [46]. This lexicon complements the prior work and provides definitions for various imaging

terms that would help standardize the use of terminology in clinical reports and research studies.

The modified Delphi method has some inherent limitations, including its reliance on expert opinion informed by evidence, possible bias in selecting experts, dependence on questionnaire design, and ability of the coordinator to effectively compile the data. We tried to minimize the dependence on questionnaire design by the coordinator by allowing the panelists to propose additional imaging terms. The reliance on the ability of the coordinator to compile data was offset by using three voting rounds and a teleconference, as well as by allowing all authors to propose any final edits to the definitions. The panel included diverse members of the SAR RCC DFP with specific expertise in renal mass imaging, representing a broad range of academic experience. We did not attempt to define terms of specific disease entities (e.g., renal cell carcinoma, angiomyolipoma); these are defined elsewhere in the literature. We only included terms used at CT and MRI as these are the most common modalities used to evaluate indeterminate renal masses. An updated lexicon that includes other modalities (e.g., ultrasound) is planned.

The Radiological Society of North American (RSNA) endorses RadLex as a lexicon designed to achieve similar goals of consistency and reproducibility for all radiology reporting [4, 47]. We did not specifically incorporate terms defined by RSNA RadLex. While there is some common ground between RadLex and this lexicon, RadLex does not contain all the terms relevant to renal mass evaluation (e.g., ‘renal mass’ or ‘growth pattern’). Additionally, some of the terms used in RadLex are not appropriate for evaluation of a renal mass (e.g., instead of ‘well-defined margin’ and ‘ill-defined margin,’ which are germane to renal mass assessment, RadLex uses ‘smooth,’ ‘lobulated’ and ‘irregular,’ which are less relevant). The panel excluded from consideration several terms because they were either confusing, non-contributory, or lacking sufficient evidence. The term ‘renal capsule’ is an anatomic term, and hence outside the scope of this lexicon. The term ‘outline’ was excluded because the panel considered ‘margin’ to represent the same feature. Terms such as calcification at MRI, central scar, and rate of enhancement were excluded because either the clinical significance of these terms was unclear, or there was not enough data to define them. Finally, terms related to texture analysis were not included because the techniques for conducting this type of analysis are not standardized, data on accuracy and reproducibility are scant, and its clinical utility is uncertain. The members of the SAR RCC DFP acknowledge that some of these terms may need to be reevaluated for potential inclusion in the lexicon in the future as new data emerge about their importance in the management of renal masses. Finally, despite our efforts to standardize terminology, there may be

persistent reader-level variability in the assessment of a renal mass [24]. We hope the specific definitions included in the lexicon will reduce this variability and take us a step closer to standardizing both radiology reports and research. Whether these goals are accomplished is a topic of further research.

In summary, this SAR-endorsed lexicon for the description of renal masses at CT and MRI has been created that attempts to address the inconsistencies and ambiguities that currently exist in clinical radiology reports and research related to renal masses. The process we used to create a renal mass lexicon may serve as a guide for the creation of lexicons in other imaging settings.

Acknowledgements We thank the Society of Abdominal Radiology Board of Directors for officially endorsing this lexicon.

Compliance with ethical standards

Conflict of interest Relevant to this work: None. Not relevant: Atul Shinagare: Consultant, Arog Pharmaceuticals, Virtualscopics. Matthew Davenport: Royalties from Wolters Kluwer. Hyesun Park: No disclosures listed. Hersh Chandarana: Research support in form of hardware and software from Siemens. Healthcare, Patent and provisional patents: MR technique GRASP and automated assessment of Image Quality with Deep Learning. Ankur M. Doshi: No disclosures listed. Ivan Pedrosa: Honorarium for a Bayer Scientific Advisory Board, co-inventor of patents with Philips Healthcare. Erick M. Remer: No disclosures listed. Nicola Schieda: No disclosures listed. Andrew Smith: President of eMASS LLC, patents pending, and President of Radiostics LLC. Raghunandan Vikram: No disclosures listed. Zhen J. Wang: Consultant, GE Healthcare; Shareholder, Nextrast, Inc. Stuart Silverman: Grant support NIH 1R21CA216796-01A1.

Appendix 1: Detailed study design

Study design: modified Delphi method

This multi-institutional, prospective, quality improvement project was exempt from IRB oversight. A modified Delphi method was used [31–34]. The Delphi technique is a structured process that relies on expert opinion and uses a series of questionnaires or ‘rounds’ to gather the required information from a selected group of experts (panelists), in this case the SAR RCC DFP. Delphi technique involves iterative, sequential, one-on-one panelist interviews. Modified Delphi technique (used in this study) involves reaching consensus by simultaneously collecting information from all panelists [48]. One radiologist member of the panel (AS, name blinded for review) served as the ‘coordinator’ who compiled the initial list of terms, prepared the questionnaires, and collected and analyzed the data. A radiology clinical fellow (HP, name blinded for review) helped create the questionnaires used in each round. To

avoid bias, the coordinator and clinical fellow did not participate in voting.

The remaining twelve radiologist members of the RCC DFP were invited to participate in the creation of the lexicon. All members are fellowship-trained abdominal radiologists with mean 13 years of experience as attending radiologist (range 5–30 years). Figure 1 outlines the process that was followed. Three rounds of questionnaires and one teleconference (after the second round) were conducted. Following the third round of questionnaires, a manuscript was created. Blinded edits to the manuscript were made and final consensus was reached. Based on prior literature, $\geq 80\%$ agreement at the end of three rounds was considered sufficient ‘consensus’ regarding the inclusion or exclusion of each term and its definition [35–38]. Individual panelist responses remained anonymous in all three rounds.

Initial selection of terms

The coordinator created an initial list of renal mass imaging features based on prior clinical and research experience and a literature search. A PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>) search through recent literature over a period of 2 years was performed from January 2016 to December 2017 using the following search string: ‘(renal or kidney) and (imaging or computed or CT or magnetic or MRI) and (features or findings).’ This yielded 60 publications. The coordinator screened the full text of these publications for imaging terms used to describe renal masses. This literature search was performed mainly to ensure that no commonly used imaging terms were missed; the actual selection of imaging terms occurred in Rounds 1 and 2 of the Delphi process during which the panelists voted on inclusion of each term and also suggested additional imaging terms if needed. The selected imaging terms were categorized into four categories: ‘basic imaging terms,’ ‘CT terms,’ ‘MRI terms’ and ‘miscellaneous terms.’

Round 1 questionnaire

The Round 1 questionnaire containing the list of the selected terms was administered using REDCap (<https://redcap.partners.org/redcap/index.php>), a secure web application for building and managing online questionnaires. Each panelist had a unique link to access their questionnaire. Each panelist was emailed up to 4 automated reminders, 1 week apart, to complete the questionnaire. All responses were submitted anonymously while blinded to the responses of the other participants.

For each term, the panelists were asked if the term should be included (Options: ‘include’ or ‘exclude’). If they selected ‘include,’ they were asked to suggest a

definition for that term in the form of free text without a word limit. During Round 1, panelists also were asked to suggest additional terms to include in the lexicon.

Round 2 questionnaire

The responses from Round 1 were analyzed by the coordinator using simple descriptive statistics. The percentage of responses for inclusion or exclusion of each term were summarized. The proposed definitions of each term were compiled to create either a single unified definition, or 2–4 alternative definitions if the content of the proposed definitions varied substantially and the coordinator was unable to coalesce them into a single definition. The summary statistics regarding inclusion and exclusion and proposed summary definition(s) were incorporated into the Round 2 questionnaire. Any rationale provided by the panelists in Round 1 to support their conflicting views was included for consideration by the other members in Round 2. In Round 2, the panelists were asked again to vote to ‘include’ or ‘exclude’ each term. For each term, if a single unified definition was suggested, the panelists were asked if they agreed with the proposed definition (Options: ‘agree’ or ‘disagree’). If they disagreed, they were required to provide an alternative definition. When more than one definition was provided, they were asked to select one of the options or to provide a new definition.

Two new terms were proposed to be added during Round 1 (‘magnetic susceptibility’ and ‘growth rate’). These were included in Round 2. Panelists were asked if these terms should be included and, if so, they were asked to suggest a definition, similar to Round 1.

The same anonymous blinded method was used to administer the Round 2 questionnaire.

Teleconference

A teleconference was conducted after the completion of Round 2 data analysis to address issues that prevented reaching consensus for terms with persistent disagreement. A summary of the discussion at the teleconference was provided to all the panelists as part of the Round 3 questionnaire.

Round 3 questionnaire

Data extracted from the Round 2 questionnaire were analyzed by the coordinator in the same fashion as the data from Round 1. A consensus definition was provided for each term. Whenever new definitions were provided, an attempt was made to reconcile these with the original proposed definition to create either a single proposed definition or a set of alternatives from which to select.

If there was 100% consensus regarding inclusion or exclusion of a particular term or definition of a term, these terms and definitions were considered ‘finalized.’ This information was provided in the Round 3 questionnaire with no further questions regarding these terms. Terms and definitions that had not met 100% consensus were included in the Round 3 questionnaire. The Round 3 questionnaire was administered to all panelists from Round 2 in the same blinded and anonymous manner as the first two questionnaire rounds. The results were summarized at the end of Round 3.

References

1. Baron RL. The radiologist as interpreter and translator. *Radiology*. 2014 Jul;272(1):4–8.
2. Breast Imaging Reporting & Data System [Internet]. [cited 2018 Jul 2]. Available from: <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/BI-Rads>
3. Liberman L, Menell JH. Breast imaging reporting and data system (BI-RADS). *Radiol Clin North Am*. 2002 May;40(3):409–30, v.
4. Langlotz CP. RadLex: a new method for indexing online educational materials. *Radiogr Rev Publ Radiol Soc N Am Inc*. 2006 Dec;26(6):1595–7.
5. Bertaud V, Lasbleiz J, Mouglin F, Burgun A, Duvauferrier R. A unified representation of findings in clinical radiology using the UMLS and DICOM. *Int J Med Inf*. 2008 Sep;77(9):621–9.
6. Filice RW, Kahn CE. Integrating an Ontology of Radiology Differential Diagnosis with ICD-10-CM, RadLex, and SNOMED CT. *J Digit Imaging*. 2019 Jan 31;
7. American College of Radiology (ACR) Illustrated Breast Imaging Reporting and Data System (BIRADS TM) American College of Radiology (ACR). Reston, VA (1998).
8. LI-RADS [Internet]. [cited 2018 Jul 2]. Available from: <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS>
9. Lung Rads [Internet]. [cited 2018 Jul 2]. Available from: <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads>
10. PI-RADS [Internet]. [cited 2018 Jul 2]. Available from: <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/PI-RADS>
11. Boland GW, Duszak R. Structured Reporting and Communication. *J Am Coll Radiol JACR*. 2015 Dec;12(12 Pt A):1286–8.
12. Schwartz LH, Panicek DM, Berk AR, Li Y, Hricak H. Improving communication of diagnostic radiology findings through structured reporting. *Radiology*. 2011 Jul;260(1):174–81.
13. Kahn CE, Heilbrun ME, Applegate KE. From guidelines to practice: how reporting templates promote the use of radiology practice guidelines. *J Am Coll Radiol JACR*. 2013 Apr;10(4):268–73.
14. Herts BR, Gandhi NS, Schneider E, Coppa CP, Mody RN, Baker ME, et al. How We Do It: Creating Consistent Structure and Content in Abdominal Radiology Report Templates. *AJR Am J Roentgenol*. 2019 Jan 8;1–7.
15. Brook OR, Brook A, Vollmer CM, Kent TS, Sanchez N, Pedrosa I. Structured reporting of multiphasic CT for pancreatic cancer: potential effect on staging and surgical planning. *Radiology*. 2015 Feb;274(2):464–72.

16. Khorasani R, Bates DW, Teeger S, Rothschild JM, Adams DF, Seltzer SE. Is terminology used effectively to convey diagnostic certainty in radiology reports? *Acad Radiol*. 2003 Jun;10(6):685–8.
17. Mityul MI, Gilcrease-Garcia B, Searleman A, Demertzis JL, Gunn AJ. Interpretive Differences Between Patients and Radiologists Regarding the Diagnostic Confidence Associated With Commonly Used Phrases in the Radiology Report. *AJR Am J Roentgenol*. 2017 Oct 12;1–4.
18. Bosmans JML, Weyler JJ, De Schepper AM, Parizel PM. The radiology report as seen by radiologists and referring clinicians: results of the COVER and ROVER surveys. *Radiology*. 2011 Apr;259(1):184–95.
19. Gunn AJ, Tuttle MC, Flores EJ, Mangano MD, Bennett SE, Sahani DV, et al. Differing Interpretations of Report Terminology Between Primary Care Physicians and Radiologists. *J Am Coll Radiol JACR*. 2016 Dec;13(12 Pt A):1525–1529.e1.
20. Rosenkrantz AB, Kiritsy M, Kim S. How “consistent” is “consistent”? A clinician-based assessment of the reliability of expressions used by radiologists to communicate diagnostic confidence. *Clin Radiol*. 2014 Jul;69(7):745–9.
21. Shinagare AB, Lacson R, Boland GW, Wang A, Silverman SG, Mayo-Smith WW, et al. Radiologist Preferences, Agreement, and Variability in Phrases Used to Convey Diagnostic Certainty in Radiology Reports. *J Am Coll Radiol JACR*. 2018 Dec 22;
22. Egbert ND, Caoili EM, Cohan RH, Davenport MS, Francis IR, Kunju LP, et al. Differentiation of papillary renal cell carcinoma subtypes on CT and MRI. *AJR Am J Roentgenol*. 2013 Aug;201(2):347–55.
23. Karlo CA, Di Paolo PL, Chaim J, Hakimi AA, Ostrovnaya I, Russo P, et al. Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations. *Radiology*. 2014 Feb;270(2):464–71.
24. Shinagare AB, Vikram R, Jaffe C, Akin O, Kirby J, Huang E, et al. Radiogenomics of clear cell renal cell carcinoma: preliminary findings of The Cancer Genome Atlas-Renal Cell Carcinoma (TCGA-RCC) Imaging Research Group. *Abdom Imaging*. 2015 Aug;40(6):1684–92.
25. Jamshidi N, Jonasch E, Zapala M, Korn RL, Aganovic L, Zhao H, et al. The Radiogenomic Risk Score: Construction of a Prognostic Quantitative, Noninvasive Image-based Molecular Assay for Renal Cell Carcinoma. *Radiology*. 2015 Oct;277(1):114–23.
26. Jinzaki M, Silverman SG, Akita H, Nagashima Y, Mikami S, Oya M. Renal angiomyolipoma: a radiological classification and update on recent developments in diagnosis and management. *Abdom Imaging*. 2014 Jun;39(3):588–604.
27. Hindman N, Ngo L, Genega EM, Melamed J, Wei J, Braza JM, et al. Angiomyolipoma with minimal fat: can it be differentiated from clear cell renal cell carcinoma by using standard MR techniques? *Radiology*. 2012 Nov;265(2):468–77.
28. Lim RS, Flood TA, McInnes MDF, Lavalley LT, Schieda N. Renal angiomyolipoma without visible fat: Can we make the diagnosis using CT and MRI? *Eur Radiol*. 2018 Feb;28(2):542–53.
29. Pokharel SS, Macura KJ, Kamel IR, Zaheer A. Current MR imaging lipid detection techniques for diagnosis of lesions in the abdomen and pelvis. *Radiogr Rev Publ Radiol Soc N Am Inc*. 2013 May;33(3):681–702.
30. Silverman SG, Pedrosa I, Ellis JH, Hindman NM, Schieda N, Smith AD, et al. Bosniak Classification of Cystic Renal Masses, Version 2019: An Update Proposal and Needs Assessment. *Radiology*. 2019;292(2):475–88.
31. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000 Oct;32(4):1008–15.
32. Powell C. The Delphi technique: myths and realities. *J Adv Nurs*. 2003 Feb;41(4):376–82.
33. Boulkedid R, Abdoul H, Loustau M, Sibony O, Albeti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One*. 2011;6(6):e20476.
34. Humphrey-Murto S, Varpio L, Wood TJ, Gonsalves C, Ufholz L-A, Mascioli K, et al. The Use of the Delphi and Other Consensus Group Methods in Medical Education Research: A Review. *Acad Med J Assoc Am Med Coll*. 2017;92(10):1491–8.
35. Sun BC, Thiruganasambandamoorthy V, Cruz JD, Consortium to Standardize ED Syncope Risk Stratification Reporting. Standardized reporting guidelines for emergency department syncope risk-stratification research. *Acad Emerg Med Off J Soc Acad Emerg Med*. 2012 Jun;19(6):694–702.
36. Berk L, Jorm AF, Kelly CM, Dodd S, Berk M. Development of guidelines for caregivers of people with bipolar disorder: a Delphi expert consensus study. *Bipolar Disord*. 2011 Sep;13(5–6):556–70.
37. Eubank BH, Mohtadi NG, Lafave MR, Wiley JP, Bois AJ, Boorman RS, et al. Using the modified Delphi method to establish clinical consensus for the diagnosis and treatment of patients with rotator cuff pathology. *BMC Med Res Methodol*. 2016 20;16:56.
38. Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. *Semin Arthritis Rheum*. 2011 Oct;41(2):95–105.
39. Schieda N, Davenport MS, Pedrosa I, Shinagare A, Chandarana H, Curci N, et al. Renal and adrenal masses containing fat at MRI: Proposed nomenclature by the society of abdominal radiology disease-focused panel on renal cell carcinoma. *J Magn Reson Imaging JMRI*. 2019 Apr;49(4):917–26.
40. Al Harbi F, Tabatabaefar L, Jewett MA, Finelli A, O’Malley M, Atri M. Enhancement Threshold of Small (< 4 cm) Solid Renal Masses on CT. *AJR Am J Roentgenol*. 2016 Mar;206(3):554–8.
41. Pallwein-Prettner L, Flöry D, Rotter CR, Pogner K, Syrè G, Fellner C, et al. Assessment and characterisation of common renal masses with CT and MRI. *Insights Imaging*. 2011 Oct;2(5):543–56.
42. Kay FU, Canvasser NE, Xi Y, Pinho DF, Costa DN, Diaz de Leon A, et al. Diagnostic Performance and Interreader Agreement of a Standardized MR Imaging Approach in the Prediction of Small Renal Mass Histology. *Radiology*. 2018;287(2):543–53.
43. Kutikov A, Uzzo RG. The R.E.N.A.L. nephrometry score: a comprehensive standardized system for quantitating renal tumor size, location and depth. *J Urol*. 2009 Sep;182(3):844–53.
44. Davenport MS, Hu EM, Smith AD, Chandarana H, Hafez K, Palapattu GS, et al. Reporting standards for the imaging-based diagnosis of renal masses on CT and MRI: a national survey of academic abdominal radiologists and urologists. *Abdom Radiol N Y*. 2017;42(4):1229–40.
45. Hu EM, Zhang A, Silverman SG, Pedrosa I, Wang ZJ, Smith AD, et al. Multi-institutional analysis of CT and MRI reports evaluating indeterminate renal masses: comparison to a national survey investigating desired report elements. *Abdom Radiol N Y*. 2018 Dec;43(12):3493–502.
46. Davenport MS, Hu EM, Zhang A, Shinagare AB, Smith AD, Pedrosa I, et al. Standardized report template for indeterminate renal masses at CT and MRI: a collaborative product of the SAR Disease-Focused Panel on Renal Cell Carcinoma. *Abdom Radiol N Y*. 2018 Dec 3;
47. RadLex [Internet]. [cited 2017 Mar 9]. Available from: <https://www.rsna.org/RadLex.aspx>
48. Ozier W. The Delphi/Modified Delphi Technique: A Consensus Approach to Information Valuation. :16.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.