Journal of Digital Imaging

Mapping LIDC, RadLex™, and Lung Nodule Image Features

Pia Opulencia,[1] David S. Channin,[2] Daniela S. Raicu,[1] and Jacob D. Furst[1]

Ideally, an image should be reported and interpreted in the same way (e.g., the same perceived likelihood of malignancy) or similarly by any two radiologists; however, as much research has demonstrated, this is not often the case. Various efforts have made an attempt at tackling the problem of reducing the variability in radiologists' interpretations of images. The Lung Image Database Consortium (LIDC) has provided a database of lung nodule images and associated radiologist ratings in an effort to provide images to aid in the analysis of computer-aided tools. Likewise, the Radiological Society of North America has developed a radiological lexicon called RadLex. As such, the goal of this paper is to investigate the feasibility of associating LIDC characteristics and terminology with RadLex terminology. If matches between LIDC characteristics and RadLex terms are found, probabilistic models based on image features may be used as decision-based rules to predict if an image or lung nodule could be characterized or classified as an associated RadLex term. The results of this study were matches for 25 (74%) out of 34 LIDC terms in RadLex. This suggests that LIDC characteristics and associated rating terminology may be better conceptualized or reduced to produce even more matches with RadLex. Ultimately, the goal is to identify and establish a more standardized rating system and terminology to reduce the subjective variability between radiologist annotations. A standardized rating system can then be utilized by future researchers to develop automatic annotation models and tools for computer-aided decision systems.

KEY WORDS: Chest CT, digital imaging, image data, image interpretation, imaging informatics, lung, radiographic image interpretation, computer-assisted, reporting, RadLex, semantic, LIDC

INTRODUCTION

An image annotation is the explanatory or descriptive information about the pixel data of an image that is generated by a human or machine observer. Much of the time, image annotation is captured as free text in a radiology report. Currently, there is little standardization of the terms used in annotations which draws us farther from the universality of semantics. The situation is further complicated when comparing human image observations to their algorithmic equivalents. That is, the way in which image pixel level data relates to the human perception of that image.

Ideally, one image should be described in the same way by multiple-trained observers. There is research[1–3] to suggest that this is not often the case (Fig. 1). Reeves et al.[1] found very high interobserver variation in lung nodule boundaries marked by radiologists. Similarly, in our previous work,[2] we showed high uncertainty and low levels of agreement between radiologist annotations when attempting to map semantic characteristics to lung nodule image content. Ochs et al.[3] showed the importance of enforcing agreement between radiologists when creating a reference standard for computer-aided diagnosis (CAD) systems.

The content and usefulness of radiological reports in diagnosis have been criticized for the amount of variability in medical terminology used in text reports.[4] To remedy this particular variability in interpretations, lexicons such as the Unified Medical Language System (UMLS), ICD-9, and SNOMED have been developed in an effort to standardize terminology.[4] However,

[1]*From the School of Computing, DePaul University, 243 S. Wabash Ave, Ste 718, Chicago, Illinois, 60604, USA.*

[2]*From the Department of Radiology, Feinberg School of Medicine, Northwestern University, 737 North Michigan Ave Ste 1600, Chicago, Illinois, 60611, USA.*

Correspondence to: Pia Opulencia, School of Computing, DePaul University, 243 S. Wabash Ave, Ste 718, Chicago, Illinois, 60604, USA; tel: +1-312-4543572; e-mail: pia@opulencia.net

doi: 10.1007/s10278-010-9285-6

Published online: 23 March 2010

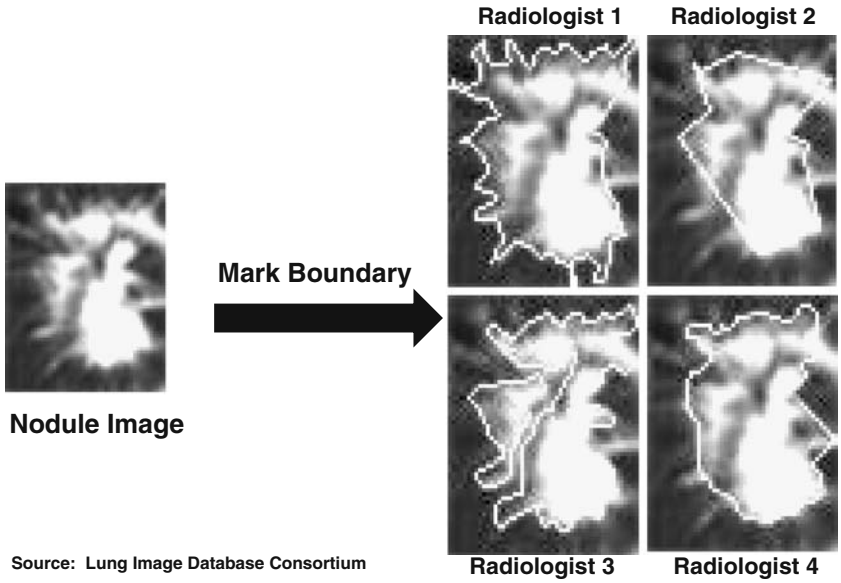

Fig 1. Example lung nodule boundaries marked by LIDC radiologists.

Langlotz and Caldwell [4] demonstrate that these lexicons have been shown to be minimally successful in capturing terms used to describe medical images in actual radiological reports. Langlotz and Caldwell [4] found that none of the lexicons achieved greater than 50% completeness for any test set of imaging terms when evaluated.

On the other hand, the Breast Imaging and Reporting Data System[5] provides a lexicon of standardized terminology, a reporting organization and assessment structure, a coding system, and a data collection structure for mammography. By specializing on a single organ and a narrow set of pathologies, McKay et al.[6] were able to demonstrate that BIRADS is an effective tool to increase interobserver objectivity, moderately high reliability between radiologists' interpretations, and moderate accuracy of interpretations. The gap between reliable subjective image interpretations is even harder to bridge with a larger set of pathologies, such as in those in the lung, and exponential development of image data.

In 2005, the Radiological Society of North America (RSNA) recognized this gap and formed a project to develop RadLex™, a radiology lexicon.[7] RadLex™ currently contains nearly 12,000 terms many of which are not found in other lexicons. It is organized in a hierarchical structure as an ontology with the primary relationship being "is_a."

In a separate endeavor, the National Cancer Institute has developed the Lung Image Database Consortium (LIDC) in an effort to provide an image database as a resource to aid in the analysis of CAD algorithms for detecting lung nodules in computed tomography (CT) scans.[8] LIDC radiologists were given the task of identifying lung nodules with size greater than or equal to 3 mm and marking their associated boundaries. The radiologists also independently rated each CT image based on nine characteristics as deemed appropriate for lung nodule diagnostic descriptors by the LIDC committee. These characteristics include: calcification, internal structure, lobulation, malignancy, margin, sphericity, spiculation, subtlety, and texture (Table 1).

The development of LIDC has led to a large amount of research based on the image sets they have provided for use. One common theme these studies share is a discussion of radiologist agreement or variability in interpreting the images.[9,10] In Ochs et al.,[3] the authors investigate the effects of radiologist agreement on the development of a "ground truth" and the subsequent impact of these effects on CAD performance. The authors of [1] investigate variability in radiologists' demarcation of nodule boundaries. Likewise, an analysis of differences in performance of radiologists' annotation methods was investigated.[11]

In comparison to radiologist ratings and annotations, which are subject to high levels of variability, lung nodule image features provide a quantitative and objective way to capture information about lung nodule images.[2,12] The difference between

Table 1. LIDC Nodule Characteristics, Definitions, and Ratings

Characteristic	Description	Ratings
Calcification	Calcification appearance in the nodule—the smaller the nodule, the more likely it must contain calcium in order to be visualized. Benignity is highly associated with central, non-central, laminated, and popcorn calcification	Popcorn Laminated Solid Non-central Central Absent
Internal Structure	Expected internal composition of the nodule	Soft tissue Fluid Fat Air
Lobulation	Whether a lobular shape is apparent from the margin or not—lobulated margin is an indication of benignity	Marked . . . None
Malignancy	Likelihood of malignancy of the nodule—malignancy is associated with large nodule size while small nodules are more likely to be benign. Most malignant nodules are noncalcified and have speculated margins.	Highly unlikely Moderately unlikely Indeterminate Moderately suspicious Highly suspicious
Margin	How well defined the margins of the nodule are	Poorly defined . . . Sharp
Sphericity	Dimensional shape of nodule in terms of its roundness	Linear . Ovoid . Round
Spiculation	Degree to which the nodule exhibits spicules, spike-like structures, along its border—spiculated margin is an indication of malignancy	Marked . . . None
Subtlety	Difficulty in detection—refers to the contrast between the lung and its surroundings	Extremely subtle Moderately subtle Fairly subtle Moderately obvious Obvious
Texture	Internal density of the nodule—texture plays an important role when attempting to segment a nodule, since part-solid and nonsolid texture can increase the difficulty of defining the nodule boundary	Nonsolid . Part-solid/mixed . Solid

radiologists' subjective, high-level interpretations and objective, low-level image features is known as a semantic gap.[21] Figure 2 provides an example of an LIDC radiologist's ratings for an image and associated low-level image features captured from the same image based on nodule boundaries marked by the radiologists.

For example, in our previous work,[2] we proposed ways to reduce the semantic gap in the medical imaging community by investigating various methods to develop computer-aided tools to be used as second readers by rating nodules based on automatically discovered image-semantic mappings. Using the LIDC dataset, we found that it is possible to develop

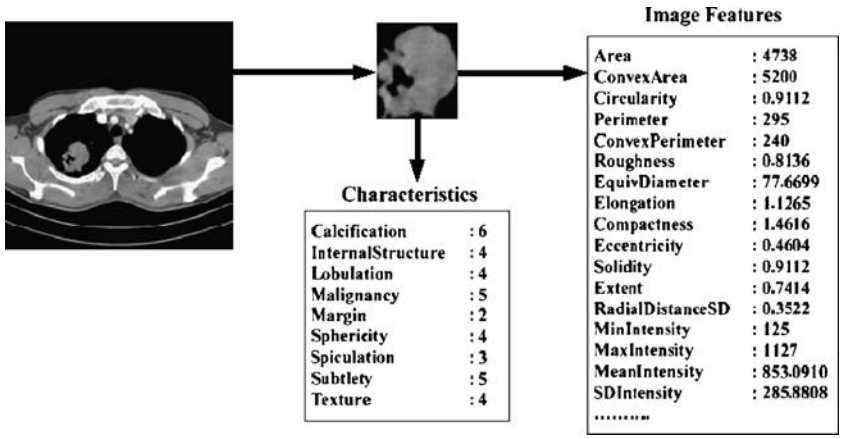

Fig 2. Example of radiologist ratings and associated features low-level image features.

probabilistic models of lung nodule image characteristics using image content. Specifically, we used low-level image features such as shape, size, intensity, and texture to develop probabilistic models for LIDC characteristics (lobulation, malignancy, margin, sphericity, spiculation, subtlety, and texture).

Using these semantic characteristics as a bridge to RadLex™'s ontology, a standardized annotation system which utilizes models learned using LIDC image features could be developed. Probabilistic models, such as those learned in[2] on the LIDC data, may then serve as decision-based rules to predict if a lung nodule can be characterized or classified with an associated RadLex™ term. As shown in Figure 3, we can then use RadLex™ terminology and predictive models based on imaging to continuously derive a standardized semantically meaningful rating system. Given the uniqueness of the publicly available LIDC dataset which provides both image data and radiologists' semantic interpretation of these data, the goal of this work is to investigate the feasibility of associating LIDC rating terminology with RadLex™ terminology. We hypothesize that the more mappings are found between image rating terminology and RadLex, the closer we are to providing a standardized system of image interpretation and diagnosis and, therefore, closer to bridging the semantic gap between image content and high-level radiologists' interpretation.

METHODS

Figure 4 presents an overview of our methodology. We searched RadLex™ for each LIDC characteristic, shown in Table 1, and their associated rating terms (for example, "sphericity" and its rating "round").[13] Each LIDC characteristic and its associated rating terms were first positioned within the RadLex™ hierarchy under related parent terms (Fig. 5). The current version of RadLex™ does not provide definitions for all of its terms, making it more difficult to accurately match terminology. In the absence of a perfect RadLex™ match, manual browsing, matching based on combined term searches (e.g., "fatty internal structure" instead of just "fatty"), and UMLS and SNOMED were used to identify terms**.** Results, therefore, are categorized in three ways: (1) exact RadLex™ matches, (2) synonymous and conceptual matches, and (3) manually searched matches. In the following sections, LIDC terms are denoted in italics while RadLex™ terms are shown in quotes. While the

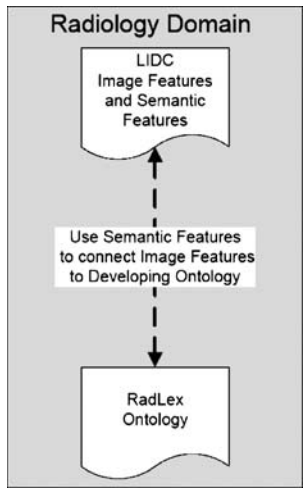

Fig 3. Linking LIDC and RadLex™.

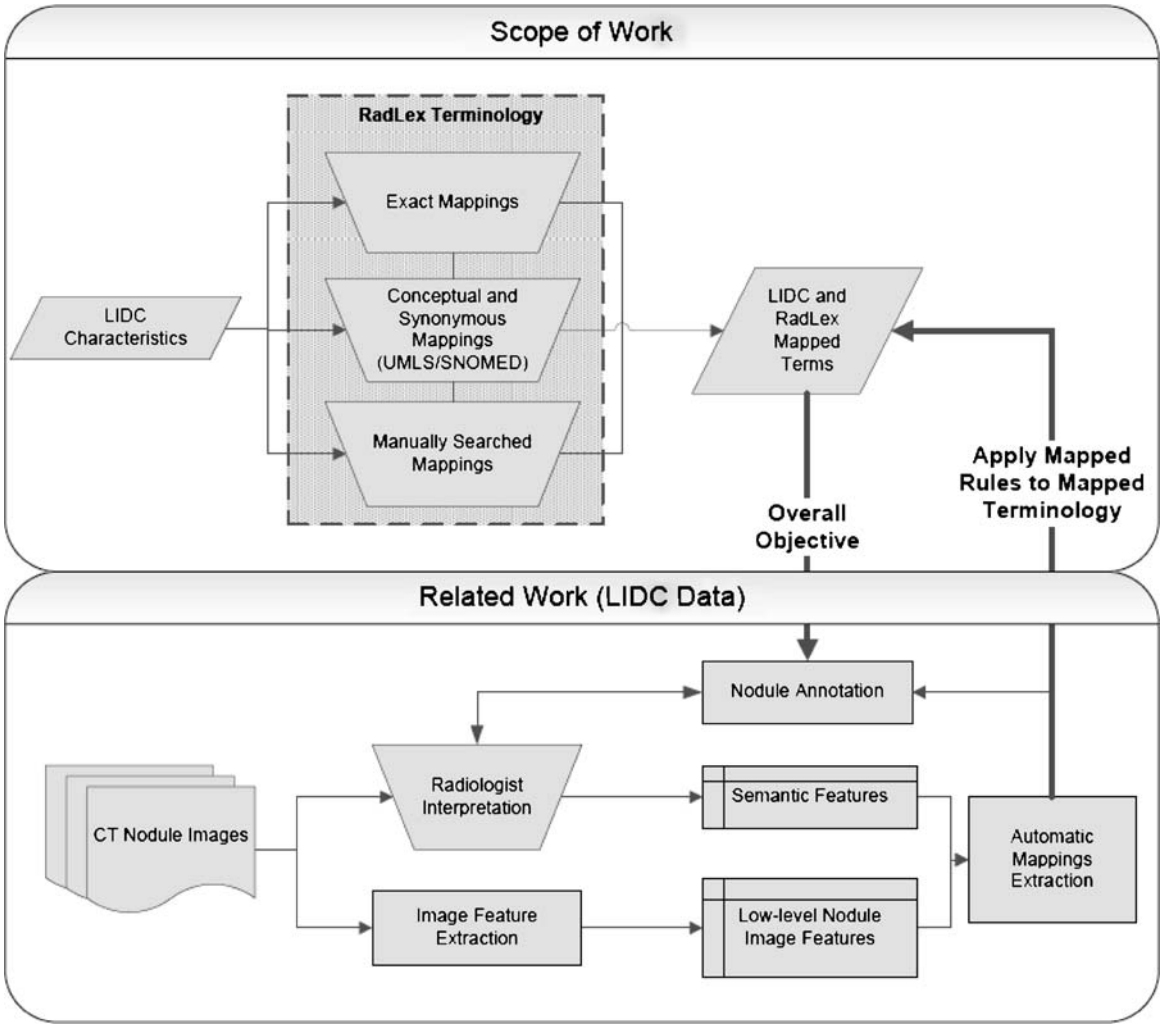

Fig 4. Diagram of proposed methodology in relation to related work.

methodology outlined above utilizes manual matching techniques with results later confirmed by an expert, automated tools do exist to find matches within RadLex. For example, WordNet could be used to automate a process to provide synonyms for terms; however, there is still the difficulty of finding terms which provide meaningful semantic matches.[14] An additional technique that could have been used is preparing an LIDC term database and to download the RadLex™ term tree. Queries could then be developed to automate the process of matching.[16] However, direct querying to RadLex™ of a relatively small number of terms did not necessitate an automated matching process. Given a larger sample of terms to match, an automated process would definitely be necessary. Likewise, in addition to an automated matching process, matches would still need to be confirmed by an expert to ensure that the meaning of the image and rating terminology are consistent.

RESULTS

Exact Matches

Sphericity

The LIDC characteristic *sphericity* does not appear within RadLex™. The LIDC *sphericity* characteristics, *linear*, *ovoid*, and *round*, however, all appear as "shapes" in RadLex™. The location of these terms in the "imaging observation char-

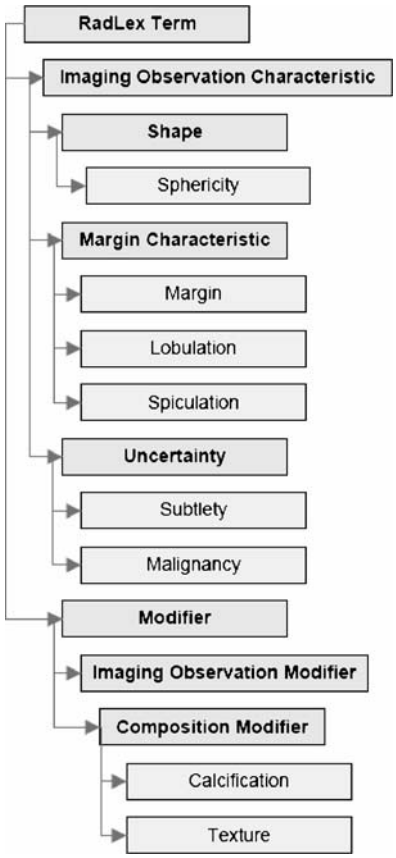

Fig 5. Mapping organization based on RadLex™ term tree.

acteristic" hierarchy suggests that they are being used harmoniously (Fig. 6).

Margin

For margin, LIDC uses an arbitrarily defined five-point scale that varies from 1, *poorly defined*, to 5, *sharp*. The term *poorly defined* is found in RadLex™ as "poorly-defined margin," both a conceptual and exact match.

The term "sharp" also appears as an exact match within RadLex™, but as a generic "morphologic characteristic." Rather, the RadLex™ term "circumscribed margin" is intended for use with respect to margins and, in fact, "sharply marginated" is present as a synonym for it.

Internal Structure

The LIDC characteristics *internal structure* contains the rating terms *soft tissue*, *fluid*, *fat*, and *air.* The term *soft tissue* is found in RadLex™ but refers to a "route of administration." Conceptually, *soft tissue* is much more closely related to the composition modifier, "solid." This should be clarified in RadLex™ and a precise definition made.

Fluid, itself, appears as an exact match, though as a generic RadLex™ "body substance." RadLex™ provides several "composition modifiers," such as "serous," "hemorrhagic," "mucinous," "proteinaceous," etc. that suggest liquid or fluid imaging characteristics, but a somewhat more generic, "fluid" or "liquid" is missing.

Fat, similarly, is found as a generic "body substance." Conceptually, the LIDC usage is much closer to the RadLex™ "composition modifier," "fat-containing." "Fatty" which appears in RadLex™ as a "morphological characteristic" also appears to be conceptually matched to LIDC's use of the term given its location in the RadLex™ tree, "imaging observation characteristic" → "morphological characteristic"→"fatty."

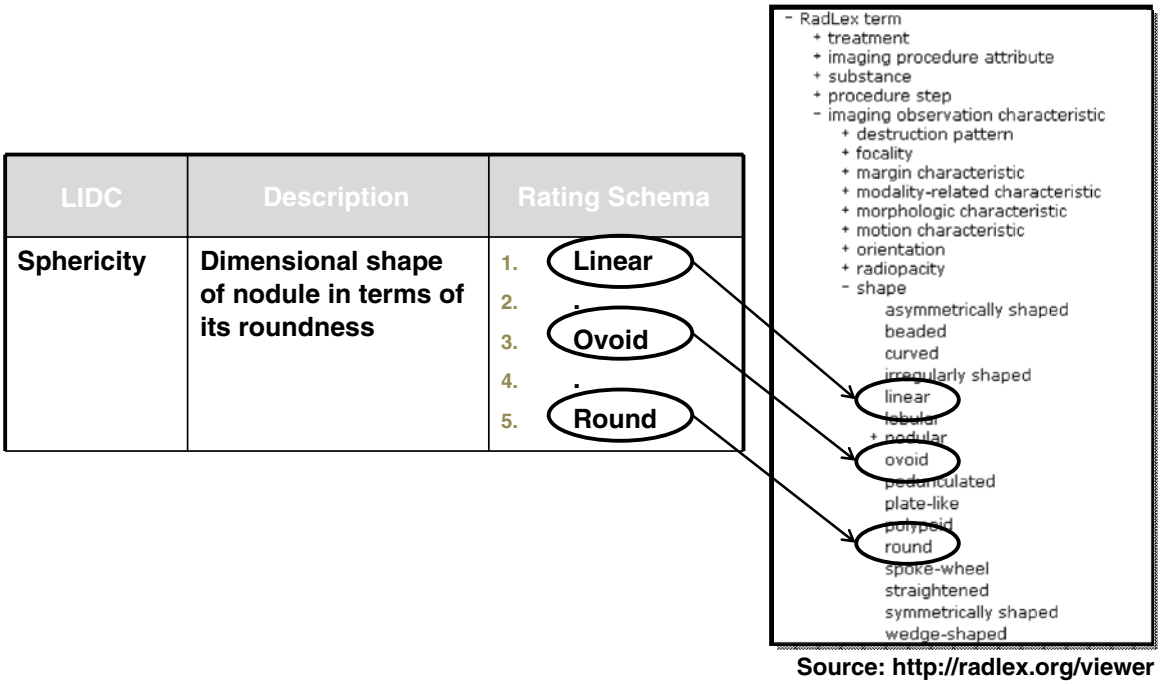

Fig 6. Example of exact matches for linear, ovoid and round.

The term "air" does not appear in RadLex™. "Air-containing" is, however, a "composition modifier" similar to "fat-containing" and most closely matches the LIDC usage.

Calcification

The LIDC characteristic *calcification* uses the rating terms *popcorn*, *laminated*, *solid*, *non-central*, *central*, and *absent*. The term *calcification* appears as an exact match in RadLex™ exactly where expected, "imaging observation"→"pathophysiologic process"→"degenerative disorder"→"deposition"→"mineral deposition"→"calcification."

An exact match for *solid* is found located in "modifier"→"imaging observation modifier"→"composition modifier"→"solid." These two terms together capture precisely the meaning intended in LIDC. Similarly, the term *central* is found in RadLex™ as a "position modifier." Again, in conjunction with the "calcification" term, this matches the LIDC meaning.

Non-central is not found in RadLex™. This is a difficult term to place as it its meaning is vague. RadLex™ contains a number of "focality" imaging observation characteristics which would have been more useful in this setting. "Scattered," "patchy," "multifocal," "focal," "diffuse," "clustered," and "coalescent" are all examples. In addition, there are "position" modifiers like "peripheral," "superficial," "superior," "inferior," "lateral," and "medial" that are related to *non-central*, however, are unrelated to the use of *non-central* as a *calcification* term.

Finally, *absent* is found in RadLex™ as a component of "definitely absent," a synonym for the RadLex™ term "definitely not present." Combined with the "calcification" term, these two terms would capture the LIDC meaning. It is interesting to note that doing so forces a choice for the calcification characteristic. The absence of the characteristic itself would indicate the lack of calcification though perhaps in a less uniform fashion.

Neither *Popcorn* nor *laminated* appear in RadLex™. *Popcorn* does appear as, "Coarse (popcorn-like) calcification" in BIRADS. When the integration of BIRADS into RadLex™ is completed in the coming year, several other terms for types of calcification will be standardized.

Laminated is more difficult. This term does not appear in BIRADS nor in RadLex™. If the intended meaning is more toward "lammelated," then the RadLex™ "shape," "plate-like" may be suitable.

Texture

Rating terms within this characteristic include *nonsolid*, *part-solid/mixed*, and *solid.* An exact match for *solid* was found located in "modifier"→"imaging observation modifier"→"composition modifier"→"solid," which belongs to both the *calcification* and *texture* characteristics. So, because *solid* has a match in RadLex™, the *texture* rating *nonsolid* would follow as a match as well. It should also be noted that since one term is shared by two different LIDC characteristics, mappings for probabilistic decision rules must specify which *solid* term it is predicting. Likewise, the term *mixed*, which belongs to the LIDC characteristic *texture*, appears in RadLex™ under "imaging observation characteristic"→ "morphological characteristic"→"mixed"; however, like other LIDC terms, we cannot say with complete confidence that this is an exact match due to its location in the RadLex™ term tree.

Lobulation and Spiculation

Both *lobulation* and *spiculation* contain the rating terms *marked* and *none.* An exact match for the characteristic terms *lobulation* and *spiculation* were not found within RadLex™. The term *marked* does appear as exact match; however, it is listed as a synonym of "severe" under "modifier"→"imaging observation modifier"→"severity modifier" in RadLex™ which suggests a conceptual match. The rating term, *none*, was not found within RadLex™.

Non-matches

Malignancy did yield an exact match within RadLex™. None of the rating terms *highly unlikely*, *moderately unlikely*, *indeterminate*, *moderately suspicious*, or *highly suspicious* appeared in RadLex™. These rating terms were searched for as a whole phrase (using both words) as well as individual words. Similarly, the rating terms for *subtlety* include *extremely subtle*, *moderately subtle*, *fairly subtle*, *moderately obvious*, and *obvious were not present.* Terms were searched for as single-word and two-word queries, and none of them were found as exact matches within RadLex™.

Synonymous and Conceptual Matches Using UMLS

According to RSNA's RadLex™ documentation available via the RadLex™ website, some of the terms used in RadLex™ appear in SNOMED. So, to identify additional search terms, synonyms were located using UMLS (2007 AC release) which includes SNOMED terminology as well. LIDC terms were then searched for within UMLS in order to identify possible synonyms. These synonyms were then used as new search terms for RadLex™. Unfortunately, few new insights or matches were made using this method of matching as illustrated by Table 2. Most UMLS-listed synonyms were not found in RadLex™ and most associated definitions did not pertain to radiology or lung nodules. As such, although some of the UMLS-listed synonyms were found in RadLex™, none of the terms (unless mapped in the previous section) could be considered as a conceptual match. For example, the terms calcification, central, soft tissue, fluid, fat, air, lobular, elevated, sharp, and spherical (denoted by parentheses in Table 2) are all synonyms of LIDC terms that were found in RadLex™, but their positions in the RadLex™ term tree suggest a conceptual mismatch. Conversely, calcified, solid, marked, moderate, border, margin, poorly (defined margin), linear, ovoid, and round, listed as UMLS synonyms, were found in RadLex™ and with conceptual matches made between UMLS definitions and positions in RadLex™. These terms were all matched in the previous section without the aid of UMLS. However, UMLS does provide definitions for some of the LIDC terms, which provide additional support for conceptual mismatches from "Exact Matches" and later, manual matching as presented in "Manually Searched Conceptual Matches."

Manually Searched Conceptual Matches

Margin Characteristics

The LIDC rating term "sharp," as discussed in the previous sections, appeared to be an exact match but was a conceptual mismatch. However, upon further investigation of RadLex™ terms under the parent term "margin characteristic," the term "circumscribed margin" appears as a child term. As shown in the RadLex™ viewer, a

Table 2. LIDC Characteristics, UMLS Definitions, and RadLex™–UMLS Matches

LIDC characteristic and possible ratings	Definitions (UMLS)	UMLS term matches and synonyms	RadLex™–UMLS matches
Calcification	Having calcium deposited on or in an anatomic structure	Calcification Calcified Calcified (qualifier value) Calcified structure Calcium deposition	(Calcification) Calcified
Popcorn		Popcorn (substance)	
Laminated		Laminar Laminar (qualifier value) Laminated	
Solid		Solid sol Solid (qualifier value) Solid—descriptor	Solid
Non-central	Central: A point or area that is approximately central within some larger region or structure (non-central would be the opposite)	Central Center Central (qualifier value)	(Central)
Absent	Not existing in a specified place at a specified time	Absent Absence of Absent (qualifier value) Negative Negative (modifier) [Ambiguous] None Not Present Ruled out	
Internal structure			
Soft tissue	Refers to muscle, fat, fibrous tissue, blood vessels, or other supporting tissue of the body	soft tissue 1A0–1A4 SOFT TISSUES Soft tissue NOS (navigational concept) Soft tissues (body structure)	(Soft Tissue)
	A continuous amorphous substance that tends to flow and to conform to the outline of its container (i.e., a liquid or a gas). Also used as an adjective to describe something	fluid - substance Fluid	
Fluid	with properties like that of a fluid.		(Fluid)
Fat			(Fat)
Air	The mixture of gases present in the earth's atmosphere consisting of oxygen, nitrogen, carbon dioxide, and small amounts of other gases	Air Air (substance) AIR PREPARATION Medical air Medical air (product)	
Lobulation	Of or relating to or resembling a lobule.	Lobular Lobular (qualifier value)	(Lobular)
Marked	A prominent or noticeable characteristic.	Marked Brisk	Marked, listed as synonym of severe
None			
Malignancy	Refers to abnormal cell activity manifested by decreased control over growth and function, causing tumor growth or spread into surrounding tissue and adverse effects to the host	Malignant—descriptor Malignant Malignant (qualifier value)	

Table 2. (continued)

LIDC characteristic and possible ratings	Definitions (UMLS)	UMLS term matches and synonyms	RadLex™–UMLS matches
Highly unlikely	Highly: An elevated level or position or degree; greater than normal in degree or intensity or amount; Unlikely: A characteristic used to qualify the adverse event as doubtfully related to the medical intervention	High Elevated High (qualifier value) Highly Unlikely Related to Intervention Doubtful Attribution Unlikely Unlikely Attribution Unlikely Related	(Elevated)
Moderately unlikely	Moderately: The quality of being within reasonable or average limits; not excessive or extreme; Unlikely: see above	Moderation Moderate Moderately (Unlikely, see above)	Moderate
Indeterminate	Cannot distinguish between two or more possible values in the current context	Indeterminate Cannot Be Determined Indeterminate (qualifier value) Undetermined Undetermined (qualifier value)	
Moderately suspicious		(Moderately, see above) SUSPICIOUS	
Highly suspicious		(Highly, see above) SUSPICIOUS	
Margin	A boundary line or the area immediately inside the boundary	Marginal Border Margin Marginal (qualifier value) Verge	Border Margin
Poorly defined		Bad Bad (qualifier value) Badly Poorly (None for defined)	Poorly, part of Poorly Defined Margin
Sharp			(Sharp)
Sphericity	A solid or hollow three-dimensional object bounded by a closed surface such that every point on the surface is equidistant from the center	Spherical shape Globular Globular shape Sphere Spherical Spherical shape (qualifier value)	(Spherical)
Linear		Linear Line Linear (qualifier value)	Linear
Ovoid		Ovoid shape Ovoid Ovoid shape (qualifier value)	Ovoid
Round	The shape of a circle. Spherical shapes are called round because their two dimensional projections are round	Round shape Circular shape Rounded Round shape (qualifier value)	Round
Spiculation	Covered with or having small needlelike structures (spicules) or fine fleshy points	Spiculate Spiculation LUNG LUL NODULE SPICULATED Spiculated lesion Lesion with spiculated margin Lesion with spiculated margin (finding)	

Table 2. (continued)

LIDC characteristic and possible ratings	Definitions (UMLS)	UMLS term matches and synonyms	RadLex™–UMLS matches
Marked	A prominent or noticeable characteristic.	Marked Brisk	Marked, listed as synonym of severe
None			
Subtlety	A subjective rating of the visibility of a lesion in a medical image, ranging from not visible to obvious	Subtlety Score Subtlety Subtlety Rating	
Extremely subtle			
Moderately subtle		Moderation Moderate Moderately (None for Subtle)	
Fairly subtle			
Moderately obvious		(see Moderately above)	
Obvious			
Texture	A measure of the variation of the intensity of a surface, quantifying properties such as smoothness, coarseness, regularity, and resiliency. The term is often used as a descriptor for the structure or organization of a tissue or organ	With texture Has texture With texture (attribute)	
Nonsolid			
Part-solid/mixed	Involving or composed of different kinds Mixed; being composed of a combination of normal and tumor cells	Mixed (qualifier value) mistura Mixed Mixed (normal and tumor)	(Mixed)
Solid		(see solid under Calcification)	Solid

Terms denoted by parenthesis indicate possible conceptual mismatches

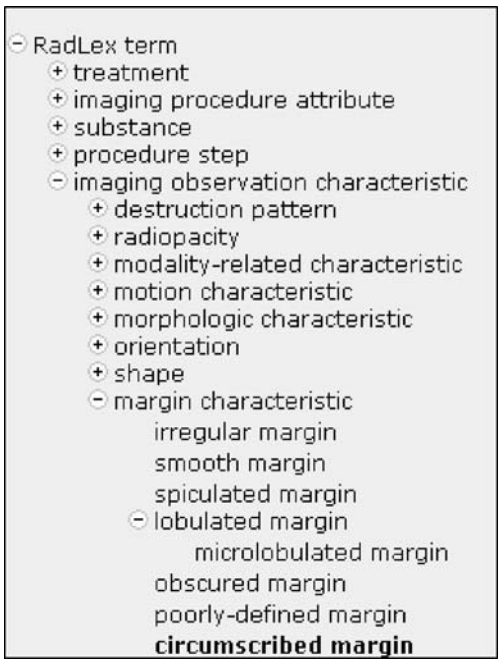

Fig 7. Screenshots of margin characteristic related matches.

synonym term for "circumscribed margin" is "sharply-marginated." The term circumscribed margin therefore provides a conceptual match to the LIDC characteristic *margin's* rating term *sharp*.

Similarly, the LIDC characteristic *lobulation* does not appear as an exact match to any RadLex™ terms. However, conceptually, *lobulation* is defined as a *margin characteristic*,[17] and

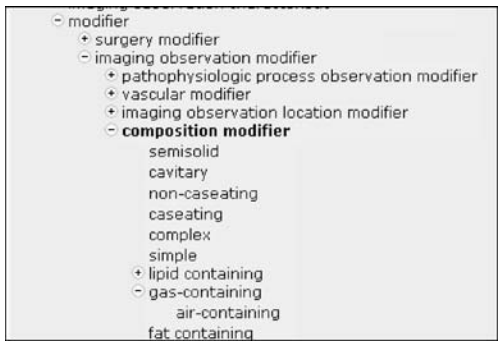

Fig 8. Screenshots of internal structure and texture related matches.

Table 3. Possible RadLex™ Matches for Subtlety and Malignancy

LIDC characteristics		RadLex™ terms
Subtlety	Malignancy	Uncertainty
1. Extremely Subtle	1. Highly unlikely	"Definitely Not Present"
2. Moderately Subtle	2. Moderately unlikely	"Almost Certainly Absent" "Probably Not Present"
3. Fairly Subtle	3. Indeterminate	"Possibly Present"
4. Moderately Obvious	4. Moderately suspicious	"Almost Certainly Present" "Probably Present"
5. Obvious	5. Highly suspicious	"Definitely Present"

when we look under the RadLex™ parent term "margin characteristic," the term lobulated margin appears with a child term "microlobulated margin." Another LIDC characteristic, *spiculation*, does not appear as an exact match in RadLex™. However, conceptually, *spiculation* is also defined as a margin characteristic as well.[18] Within the RadLex™ term browser, the term spiculated margin appears under the RadLex™ parent term "margin characteristic." Figure 7 provides screen shots of examples the conceptual matches for *sharp*, *lobulation*, and *spiculation* in RadLex™.

Specifically, Figure 7 illustrates the importance of term relationships within a RadLex™ header. It further supports matching methods described in "Exact Matches" and "Synonymous and Conceptual Matches Using UMLS" in that RadLex™ child terms located within a parent header can either be identified as conceptually related or conceptually unrelated to a certain LIDC term.

Internal Structure, Calcification, and Texture Revisited

Terms listed under "composition modifier" appear to be conceptually matched to LIDC characteristics related to both the *internal structure* and *calcification* of a nodule. Specifically, a conceptual match for *fat* was found under "modifier"→"imaging observation modifier"→"composition modifier"→"fat-containing." In the same section of RadLex™, as seen in Figure 8, a match for *air* can be found under "gas-containing"→"air-containing." A conceptual match for *fluid* was found under the same header under the term "serous." Serous is defined as "of thin watery constitution" (Dictionary.com, retrieved March 25, 2008). Similarly, the *texture* rating term *part-solid/mixed* has a conceptual match under "modifier"→"imaging observation modifier"→"composition modifier"→"semisolid." Likewise, as mentioned in "Margin," *calcification* was found in RadLex™ as an exact match; however, its location within the term tree suggests that there might be a better match. As such, a better match for *calcification is* the term "calcified" which appears under "modifier"→"imaging observation modifier"→"composition modifier"→"calcified."

Malignancy and Subtlety

While no exact matches for ratings terms in *malignancy* and *subtlety* were found in RadLex™, a match for the *calcification* rating term *absent* leads to the discovery of the RadLex™ header "uncertainty." Other RadLex™ terms listed under this header include "definitely not present," "almost certainly absent," "probably not present," "possibly present," "almost certainly present," "probably present," and "definitely present." These terms suggest a possible conceptual match with ratings terms for both malignancy and subtlety. A proposed matching schema is presented in Table 3. Using RadLex™ terms as standardized rating terms for both subtlety and malignancy may provide a way for decision rules to be more easily applied; however, it may not convey the intended interpretation exactly.

Table 4. Example of Matched Term with Decision Rule

LIDC	RadLex™ matching term	Possible decision rule
Subtlety		
1. Extremely subtle	Definitely not present—shared	**Subtlety1_Rule1** *{Pr(1)=1.00}* (minorAxisLength<=0.14683) and (maxIntensity<=0.226443)

DISCUSSION

After performing all three methods for matching LIDC terminology to RadLex™ terms, 74% of the terms (25 out of 34) were found in RadLex™. Thus, the mapped terminology can now be matched with probabilistic models developed using an LIDC dataset which can now serve as decision-based rules to predict if a lung nodule can be characterized or classified with an associated RadLex™ term. Table 4 contains an example of a decision rule mappings to *subtlety*. The decision rule contains the image features maxIntensity (a gray-level intensity feature) and minorAxisLength (a size feature) and applies the following logic to assign a rating: if (minorAxisLength≤0.15) and (maxIntensity≤0.23), then the associated *subtlety* rating is 1 or *extremely subtle* with a 100% confidence. The particular decision rule presented in Table 4 was based on our previous research.[2]

Table 5 summarizes our results; specifically, it contains all LIDC terms with matches in RadLex™. Items denoted as shared indicated that the same RadLex™ term is shared by two or more LIDC terms. So, any decision rules applied to the RadLex™ term must specify which LIDC term it belongs to. Similarly, items denoted as "opposite of" indicate that the matched term for any particular rating (e.g., *None*) is simply the opposite of the rating term on the other end of the scale (e.g., *Lobulated Margin*). So, in lieu of the fact that there is no match for the LIDC rating term *none*, a nodule with the predicted annotation as *none* using *lobulation rating 5* decision rules would be interpreted as having the opposite meaning of *lobulated margin* or would be interpreted as (*not*) *lobulated margin*.

CONCLUSION

The results of this study identified matches for 25 out of 34 LIDC terms in RadLex™. With most of the LIDC terms now mapped to RadLex™, predictive rules for annotation were applied as learned by our previous research.[2] With these rules, a CT image of a lung nodule can be interpreted by a computer, annotated, and further verified by a radiologist using similar descriptors. On a larger scale, automatic lung nodule annotation based on low-level image features can narrow the semantic gap. Predictive rules for annotating

Table 5. Summary of Matched Terms

LIDC	RadLex™ matching term
Sphericity	
1. Linear	Linear
2. .	
3. Ovoid	Ovoid
4. .	
5. Round	Round
Margin	
1. Poorly Defined	Poorly defined margin
2. .	
3. .	
4. .	
5. Sharp	Circumscribed Margin (Synonym: Sharply-Marginated)
Spiculation	
1. Marked	Spiculated Margin
2. .	
3. .	
4. .	
5. None	(opposite of) Spiculated Margin
Lobulation	
1. Marked	Lobulated Margin
2. .	
3. .	
4. .	
5. None	(opposite of) Lobulated Margin
Internal Structure	
1. Soft Tissue	
2. Fluid	Serous
3. Fat	Fat-Containing
4. Air	Air-Containing
Malignancy	
1. Highly Unlikely	Definitely Not Present - shared
2. Moderately Unlikely	Probably Not Present
3. Indeterminate	Possibly Present—shared
4. Moderately Suspicious	Probably Present
5. Highly Suspicious	Definitely Present—shared
Subtlety	
1. Extremely Subtle	Definitely Not Present—shared
2. Moderately Subtle	Almost Certainly Absent
3. Fairly Subtle	Possibly Present—shared
4. Moderately Obvious	Almost Certainly Present
5. Obvious	Definitely Present—shared
Texture	
1. Non-Solid	(opposite of) Solid
2..	
3. Part-Solid/Mixed	Semisolid
4. .	
5. Solid	Solid—shared
Calcification	
1. Popcorn	
2. Laminated	
3. Solid	Solid—shared
4. Non-Central	
5. Central	Central
6. Absent	Absent

images learned using the LIDC dataset and terminology were mapped to RadLex™ terminology in an effort to reduce subjective variability of radiologist image interpretations. Future work may include re-annotating lung nodule images based on a modified LIDC terminology which utilizes RadLex™ terms. Work can also be done using RadLex™ terminology directly in an effort to develop a new, radiologist-annotated image database by which future researchers may develop predictive rules and apply them to a larger, more standardized CAD system. Likewise, additional work may include an implementation of clustering analysis or some other unsupervised learning technique to uncover new meanings, labels, or characteristics for groups of nodules.

Not all LIDC terms, however, were found to have matches in RadLex™, and many of the mapped terms were not found as exact matches. The fewest number of matches was for *calcification*, suggesting that the terminology used by LIDC may be inconsistent with other terms found in existing lexicons. This suggests that there are opportunities to better conceptualize, define, or reduce LIDC characteristics and associated rating terminology in an effort to produce even more matches with RadLex™. It is also important to note that even with the use of UMLS and SNOMED to help define LIDC terminology, no new additional matches were discovered. As shown in Table 2, numerous synonyms for each LIDC characteristic were provided by SNOMED; however, when these synonyms were searched for in RadLex™, no additional matches were found in comparison to direct querying of RadLex™ of the original LIDC terms. As such, while there are several lexicons, ontologies, and image reporting initiatives that exist and are being developed,[15,16] a single concise lexicon and ontology would greatly advance radiology reporting consistency and accuracy.[19,20]

By utilizing LIDC semantic characteristics as a bridge to RadLex™'s ontology, a standardized annotation system based on models learned using LIDC image features could be developed. The universality of radiological concepts and terminology is what this research aims to capture, in an effort to reduce the semantic gap between image content and high level radiologists' interpretation, and ultimately to provide a standardized system of image interpretation and diagnosis.

REFERENCES

1. Reeves AP, Biancardi AM, Apanasovich TV, Meyer CR, MacMahon H, van Beek EJR, Kazerooni EA, et al: The Lung Image Database Consortium (LIDC): pulmonary nodule measurements, the variation, and the difference between different size metrics. Proc SPIE Int Soc Opt Eng 8:65140J.1–65140J.8, 2007, doi:10.1117/12.713672

2. Raicu DS, Varutbangkul E, Furst JD, Armato III, SG: Modeling semantics from image data: opportunities from LIDC. IJBET 2:1–22, 2008

3. Ochs R, Kim HJ, Angel E, Panknin C, McNitt-Gray M, Brown M: Forming a reference standard from LIDC data: impact of reader agreement on reported CAD performance, Medical Imaging 2007: Computer-Aided Diagnosis. Proc SPIE Int Soc Opt Eng 6514:65142A, 2007, doi:10.1117/12.707916

4. Langlotz CP, Caldwell S: The completeness of existing lexicons for representing radiology report information. J Digit Imaging 15:201–205, 2001, doi:10.1007/s10278-002-5046-5

5. Eberl MM, Fox CH, Edge SB, Carter CA, Mahoney MC: BI-RADS Classification for management of abnormal mammograms. JABFM 19:161–164

6. McKay C, Hart CL, Erbacher G: Objectivity and accuracy of mammogram interpretation using the BI-RADS final assessment categories in 40- to 49- year old women. JAOA 100 (10):615–620, 2000

7. Langlotz CP: RadLex™: A new method for indexing online educational materials. RadioGraphics doi:10.1148/rg.266065168, September 14, 2006

8. Armato III, SG, McLennan G, McNitt-Gray MF, Meyer CR, Yankelevitz D, Aberle DR, Henschke CI, Hoffman EA, Kazerooni EA, MacMahon H, Reeves AP, Croft BY, Clarke LP: Lung image database consortium: Developing a resource for the medical imaging research community. Radiology 232:739–748, 2004, doi:10.1148/radiol.2323032035

9. Armato III, SG, McNitt-Gray MF, Reeves AP, Meyer CR, McLennan G, Aberle DR, Kazerooni EA, MacMahon H, van Beek EJR, Yankelevitz D, Hoffman EA, Henschke CI, Roberts RY, Brown MS, Engelmann RM, Pais RC, Piker CW, Qing D, Kocherginsky M, Croft BY, Clarke LP: The Lung Image Database Consortium (LIDC): An evaluation of radiologist variability in the identification of lung nodules on CT scans. Acad Radiol 14:1409–1421, 2007

10. Armato III, SG, Roberts RY, Kocherginsky M, Aberle DR, Kazerooni EA, MacMahon H, van Beek EJR, Yankelevitz DF, McLennan G, McNitt-Gray MF, Meyer CR, Reeves AP, Caligiuri P, Quint LE, Sundaram B, Croft BY, Clarke LP: Assessment of radiologist performance in the detection of lung nodules: Dependence on the definition of "truth.". Acad Radiol 16 (1):28–38, 2008

11. Meyer CR, Johnson TD, McLennan G, Aberle DR, Kazerooni EA, MacMahon H, Mullan BF, Yankelevitz DF, van Beek EJR, Armato III, SG, McNitt-Gray MF, Reeves AP, Gur D, Henschke CI, Hoffman EA, Bland PH, Laderach G, Pais R, Qing D, Piker C, Guo J, Starkey A, Max D, Croft BY, Clarke LP: Evaluation of lung MDCT nodule annotation across radiologists and methods. Acad Radiol 13:1254–1265, 2006

12. Horsthemke W, Varutbangkul E, Raicu D, Furst J: Predictive data mining for lung nodule interpretation. In

Proceedings of the Seventh IEEE international Conference on Data Mining Workshops (October 28–31, 2007). ICDMW. IEEE Computer Society, Washington, DC, 157–162. doi:http://dx.doi.org/10.1109/ICDMW.2007.74, 2007.

13. McNitt-Gray MF, Armato III, SG, Meyer CR, Reeves AP, McLennan G, Pais R, Freymann J, Brown MS, Engelmann RM, Bland PH, Laderach GE, Piker C, Guo J, Towfic Z, Qing DP, Yankelevitz DF, Aberle DR, van Beek EJR, MacMahon H, Kazerooni EA, Croft BY, Clarke LP: The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation. Acad Radiol 14:1464–1474, 2007

14. Fellbaum C: WordNet: An Electronic Lexical Database, 1998

15. Channin DS: The caBIG Annotation and Image Markup Development Project, Chicago: Radiological Society of North America Annual Meeting and Scientific Assembly, 2006

16. Rubin D, Channin DS, Mongolowat P, et al: LIDC Conversion to AIM, February 13, 2009. Last Accessed March 12, 2009. Available at: https://wiki.nci.nih.gov/display/Imaging/LIDC+Conversion+to+AIM

17. Sluimer I, Schilhan A, Prokop M, Ginneken B: Computer analysis of computed tomography scans of the lung: a survey. IEEE Trans Med Imaging 25:385–405, 2006, doi:10.1109/TMI.2005.862753

18. Zerhouni EA, Stitik FP, Siegelman SS, Nadidich DP, Sagel SS, et al: CT of the pulmonary nodule: a cooperative study. Radiology 160:319–327, 1986

19. Vanel D: The American College of Radiology (ACR) Breast Imaging and Reporting Data System (BI-RADS™): A step towards a universal radiological language? Eur J Radiol 61 (2):183–183, 2007

20. Sistrom CL, Langlotz CP: A framework for improving radiology reporting. J Am Coll Radiol 2(2):159–167, 2005

21. Deserno TM, Antani S, Long LR: Exploring access to literature using content-based image retrieval. SPIE Medical Imaging 6516, 2007