Published in final edited form as:

*Acad Radiol*. 2020 January ; 27(1): 76–81. doi:10.1016/j.acra.2019.09.011.

# Medical Image Analysis: Human and Machine

**Robert Nick Bryan, MD, PhD**, **Christos Davatzikos, PhD**, **Edward H. Herskovits, MD, PhD**, **Suyash Mohan, MD**, **Jeffrey D. Rudie, MD, PhD**, **Andreas M. Rauschecker, MD, PhD**

Department of Diagnostic Medicine, University of Texas at Austin Dell Medical School, 1601 Trinity St., Bldg. B, Stop Z0300, Austin, TX 78712 (R.N.B.); Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, Pennsylvania (C.D.); Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland, Maryland (E.H.H.); Department of Radiology, Division of Neuroradiology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania (S.M.); Department of Radiology and Biomedical Imaging, University of California, San Francisco, California (J.D.R., A.M.R.).

### Keywords

Neuroimaging; Computer; Image analysis

## INTRODUCTION

**I**mages play a critical role in science and medicine. Until recently, analysis of scientific images, specifically medical images, was an exclusively human task. With the evolution from analog to digital image data and the development of sophisticated computer algorithms, machines (computers) as well as humans can now analyze these images. Though different in many ways, these alternative means of analysis share much in common: we can model the newer computational methods on the older human-observer approaches. In fact, our Perspective holds that radiologists and computers interpret medical images in a similar fashion.

Critical to understanding image analysis, whether by human or machine, is an appreciation of what an image is. The universe, including all and any part therein, can be defined as:

$$\mathrm{U} = (m, E)(x, y, z)(\mathrm{t})$$

Or, more qualitatively: the universe is mass and energy distributed within a three-dimensional space, varying with time (1). Scientific observations, whether encoded as numerical measurements or categorical descriptors, reflect information about one or more of these three intrinsic domains of nature. Measurements of mass and energy (m, E) are often called signals. In medical imaging, essentially all our signals reflect measurements of energy. An image can be defined as a rendering of spatially and temporally defined signal measurements, or:

**Address correspondence to:** R.N.B. Nick.Bryan@austin.utexas.edu.

$$I = f(m, E)(x, y, z)(t)$$

Note the parallelism between what the universe is and how it is reflected by an image. In the context of scientific observations, an image is the most complete depiction of observations of nature— or in the case of medicine, the patient and his/her disease. Images and images alone include explicit spatial information, information that is intrinsic and critical to the understanding of most objects of medical interest.

## HUMAN IMAGE ANALYSIS

Pattern recognition is involved in many, if not most, human decisions. Human image analysis is based upon pattern recognition (2). In medicine, radiologists use pattern recognition when making a diagnosis. It is the heart of the matter. Pattern recognition has two components: pattern learning and pattern matching. Learning is a training or educational process; radiology trainees are taught the criteria of a normal chest X-ray examination and observe hundreds of normal examinations, eventually establishing a mental pattern of "normal." Matching involves decision-making; when an unknown chest film is presented for interpretation, the radiologist compares this unknown pattern to their "normal" pattern and makes a decision as to whether or not the case is normal, or by exclusion, abnormal.

One of the most striking aspects of human image analysis is how our visual system deconstructs image data in the central nervous system. The only place in the human head where there is anything resembling a coherent pattern of what we perceive as an image is on the surface of the retina. Even at this early point in the human visual process, the image data are separated into color versus light intensity pathways, and other aspects of the incoming image are emphasized and/or suppressed. The deconstruction of image data proceeds through the primary visual cortex into secondary, and higher, visual cortices where different components of image data, particularly those related to signal (brightness), space (shape), and time (change), are processed in distinct and widely separated anatomic regions of the brain (3, 4). The anatomic substrate of these cortical patches that process image data is the six layered, cerebral cortex columnar stacks of neurons that make up local neural networks (5).

The deconstruction of image data in the human brain— "what happens where"—is relatively well understood. However, the structures and processes involved in the *reintegration* of these now disparate data into a coherent pattern of what we perceive remains a mystery. Regardless of how the brain creates a perceived image, the knowledge that it does so by initially deconstructing image data and processing these different data elements by separate anatomic and physiological pathways provides important clues as to how images are analyzed by humans.

## FINDINGS = OBSERVED KEY FEATURES

In the process of deconstructing image data, the human brain extracts key, or salient, features by separate mechanisms and pathways. These key features (KFs) reveal the most important

information. While their number and nature are not fully understood, it is clear that KFs include signal, spatial, and temporal information. They are separately extracted and analyzed with the goal of defining dominant patterns in the image that can be compared to previously learned KF patterns in the process of diagnostic decision-making.

Given that analyzing the patterns of KFs of an image is fundamental to human image interpretation, an obvious step in the interpretive process is to define the KFs to be extracted and the patterns thereof learned. This is the learning part of pattern recognition and is a function of image data itself, empirical experience, and task definition. KFs must be contained in the image data, extractable by an observer, and relevant to the decision-making process. Since image data consist of signal, spatial, and temporal information, KFs will likewise reflect one or more of these elements. To extract a KF, a human observer has to be able to see the feature on an image. In a black and white image, a KF cannot be color. Ideally, KFs are easy to see and report, i.e., extract.

A KF must make a significant contribution to some decision based on the image data. Since the ultimate performance metric of medical image interpretation is diagnostic accuracy, KFs of medical images must individually contribute to correct diagnoses. The empirical correlation of image features with specific diagnosis by trial and error observations has been the traditional way to identify KFs on the basis of their contribution to diagnosis. Observed KFs provide the content for the “Findings” section of a radiology report. For brevity and convenience, we will focus on signal and spatial KFs.

Medical diagnosis in general and radiological diagnosis in particular is organ based. The brain is analyzed as a separate organ from the “head and neck,” which must be separately analyzed and reported. Every organ has its unique set of KFs and disease patterns. The first step in the determination of “normal” requires the learning of a pattern, which in this case consists of the signal and spatial KFs of a normal brain.

For medical images, the signal measured by the imaging device is usually invisible to humans, and therefore the detected signal must be encoded as visible light, most commonly as the relative brightness of pixels or voxels. In general, the greater the magnitude of the signal detected by the imaging device, the brighter the depiction of the corresponding voxel in the image. Once again, the first step in image analysis is to extract a KF from the image, in this case relative voxel brightness. An individual image’s signal pattern is compared to a learned normal pattern. If the signal pattern of an unknown case does not match the normal pattern, one or more parts of the diagnostic image must be brighter or darker than the anatomically corresponding normal tissue.

The specific nature of an abnormal KF is summarized in the Findings section of the radiology report, preferably using very simple descriptors, such as “Increased” or “Decreased” signal intensity (SI). To reach a specific diagnosis, signal KFs for normal and abnormal tissues must be evaluated, though usually only abnormal KFs are reported. Signal KFs are modality specific. The SIs of different tissues are unique to that modality. For each different signal measured, there is usually a modality specific name (with X-ray images, for

example, radiodensity is a name commonly applied to SI, with relative intensities described as increased or decreased).

Specific objects within images are initially identified and subsequently characterized on the basis of their signal characteristics. The more unique the signal related to an object, the simpler this task. For example, ventricles and subarachnoid spaces consist of cerebrospinal fluid, which has relatively distinctive SI on computed tomography (CT) and magnetic resonance imaging (MRI). Other than being consistent with the signal of the object of interest (i.e., cerebrospinal fluid), SI is irrelevant to the evaluation of the spatial features of that object. With minimal training, most physicians can easily “extract”, i.e., see and distinguish signal KFs on the basis of relative visual brightness.

The second component of the Findings section of a radiology report relates specifically to spatial components of image data. Spatial analysis is geometric in nature and commonly uses geometric descriptors for spatial KFs. The most important spatial KFs are number, size, shape and anatomic location. A prerequisite for the evaluation of these spatial attributes is identification of the object to which these descriptors will be applied, beginning with the organ of interest.

In the case of the brain, we uniquely use surrogate structures, the ventricles and subarachnoid spaces, to evaluate this particular organ’s spatial properties. Due to the fixed nature of the adult skull, ventricles and subarachnoid spaces provide an individually normalized metric of an individual’s brain size, shape, and position. Fortunately, the ventricles and subarachnoid spaces can easily be observed on CT, MRI, or ultrasound and their spatial attributes easily learned by instruction and repetitive observations of normal examinations.

The second step of pattern recognition—pattern matching—is completely dependent on the first step of pattern recognition—pattern learning. Matching is the operative decision-making step of pattern recognition. In terms of ventricles and subarachnoid spaces, the most fundamental spatial pattern discriminator is size, whether or not the ventricles are abnormally large or small. If the ventricles and/or subarachnoid spaces are enlarged, the differential diagnoses might include hydrocephalus or cerebral atrophy. If they are abnormally small, mass effect is suggested, and the differential diagnosis might include cerebral edema, tumor, or other space occupying lesion. In any case, a KF extracted from any brain scan is the spatial pattern of the ventricles and subarachnoid spaces, this specific pattern is matched against a learned, experience-based, normal pattern, and a decision of normal or abnormal is made.

When reporting image features, humans tend to use categorical classification systems rather than numeric systems. Humans will readily, though not always reliably, classify a light source as relatively bright or dark, but only reluctantly attempt to estimate the brightness in lumens or candelas. Humans are not good at generating numbers from a chest film, but they are very good at classifying it as normal or abnormal. If quantitative image measurements are required, radiologists bring additional measurement tools to bear, like a ruler to measure the diameter of a tumor, or a computer to calculate its volume. If pushed for a broader, more

dynamic reporting range, a radiologist may incorporate a qualitative modifier, such as "marked," to an abnormal KF description to indicate the degree of abnormality.

Interestingly, and of practical importance, human observers tend to report psychophysiological observations using a scale of no more than seven. This phenomenon is well documented in George Miller's paper, *The Magical Number 7* (6). A comparative scale of seven is reflected in the daily use of such adjective groupings as "mild, moderate, severe"; "possible, probable, definite"; "minimal, moderate, marked." If an image feature has the possibility of being normal, increased, or decreased, with three degrees of abnormality in each direction, the feature can be described with a scale of seven. While there are other human observer scales, feature rating scales of from two to seven generally suffice and reflect well documented behavior of radiologists (7).

Based on the concept of extracting a limited number of KFs and reporting them with a descriptive scale of limited dynamic range, it is relatively straightforward to develop a highly structured report-generating tool applicable to diagnostic imaging studies. The relative intensity of each imaging modality's detected signal is a KF, potentially reflecting normal or pathological tissue. An accompanying spatial KF of any abnormal signal is its anatomic location. A spatial KF of brain images is the size of the ventricles and subarachnoid spaces, which reflect the presence or absence of mass effect and/or atrophy.

## IMPRESSION = INFERRED DIFFERENTIAL DIAGNOSIS

Medical diagnosis is based upon the concept of differential diagnoses, which consist of a list of diseases with similar image findings. A radiographic differential diagnoses is the result of the logically consistent matching of KFs extracted from a medical image to specific diagnosis. KFs are extracted from medical images, summarized by structured descriptive findings as previously described, and a differential diagnostic list consistent with the pattern of extracted features is inferred. This inferential form of pattern matching for differential diagnosis is reflected in such publications as *Gamuts of Differential Diagnosis* and *StatDx* (8, 9). These diagnostic tools consist of a list of diseases and a set of matching image KFs.

Differential diagnosis, therefore, is another pattern recognition process based upon the matching of extracted KF patterns to specific diseases. A complete radiographic report incorporates a list of observed KFs summarized in the FINDINGSFINDINGS and a differential diagnosis in the IMPRESSION, which was inferred from the KFs. A normal x-ray CT report might be:

### Findings

- There are no areas of abnormal radiodensity. (Signal features encoded as relative light intensity)
- The ventricles and subarachnoid spaces are normal as to size, shape, and position. (Spatial features of the organ of interest, the brain)
- There are no craniofacial abnormalities. (Signal/spatial features of another organ)
- There is no change from the previous exam. (Temporal feature)

**Impression**

- Normal examination of the head. (Logical inference)

For an abnormal report, one or more of the KF statements must be modified and the Impression must include one or more inferred diseases.

**Findings**

- There is increased radiodensity in the right basal ganglia.
- The frontal horn of the right lateral ventricle is abnormally small (compressed).
- There are no craniofacial abnormalities.
- The lesion was not evident on the previous exam.

**Impression**

- Acute intracerebral hemorrhage.

The list of useful KFs is limited by the nature of signal and spatial data and is, we believe, relatively short. While human inference mechanisms are not fully understood, the final diagnostic impression probably reflects rule-based or Bayesian processes, the latter of which deal better with the high degree of uncertainty in medicine and take better advantage of prior knowledge, such as prevalence of disease in a practice (8).

Less experienced radiologists and radiology trainees typically perform image analysis as outlined above, tediously learning and matching normal and abnormal signal and spatial patterns, consciously extracting KFs, and then deducing the best matches between the observed KFs and memorized KF patterns of specific diseases. This linear intellectual process is an example of “thinking slow,” a cognitive process described by Kahneman (10). However, when a radiologist is fully trained and has sufficient experience, he/she switches from this cognitive mental process to the much quicker “thinking fast,” heuristic mode of most professional practitioners in most fields. Most pattern matching tasks take less than a second to complete. A skilled radiologist makes the normal/abnormal diagnosis of a chest image in less than one second (11).

In his book *Outliers*, Malcom Gladwell famously concluded that 10,000 hours of training are mandatory to function as a professional (12). The specific number has been challenged, of course, but it appropriately emphasizes the fact that professionals’ function differently than amateurs. They think fast, and, often, accurately. To achieve success at this level, the professional needs to have seen and performed the relevant task thousands of times—exactly how many thousand, who knows. The neuropsychological processes underlying these “slow” and “fast” mental processes are not clear, but it is hypothesized that higher order pattern matching processes become encoded in brain structure and eventually allow the “ah hah” identification of an “Aunt Minnie” brain stem cavernoma in a fraction of a second on a T1-weighted MRI image.

However, humans working in this mode do make mistakes related to well-known biases, including: availability (recent cases seen), representativeness (patterns learned), and

anchoring (prevalence) (13). Other psychophysical factors such as mood and fatigue can also affect this process. Slower, cognitive thinking does not have the same faults and biases. The two types of decision-making are complementary and often combined, as in the case of a radiologist interpreting a case of a rare disease that they have not seen or a case with a disease having a more variable KF pattern.

## COMPUTER IMAGE ANALYSIS

Whereas humans can analyze analog or digital images, computers can operate only on digital or digitized images, both types of which can be defined as before:

$$I = f((m, E)(x, y, z)(t))$$

Therefore, computers face the same basic image analysis problem as humans and can perform this task similarly. As with human observers, computers can be programmed to deconstruct an image in terms of signal, spatial, and temporal content. It is relatively trivial to develop and implement algorithms that extract the same image KFs from digital data that radiologists extract from analog or digital data. Computers can be trained with pattern recognition techniques to match image KFs with normal and/or disease feature patterns in order to formulate a differential diagnosis.

A significant difference between human and computer image analysis is the relative strength in classifying versus quantifying image features. Humans are very adept at classifying observations but can quantify them only crudely. In contrast, quantitative analysis of scientific measurements is the traditional forte of computers. Until recently, computers tended to use linear algebraic algorithms for image analysis (14), but with the advent of inexpensive graphics processing unit hardware and neural network algorithms, classification techniques are being widely implemented (15). Each approach has different strengths and weaknesses for specific applications, but combinations of the two will offer the best solutions for the diverse needs of the clinic.

To illustrate these two computational options for image analysis, let us take the task of extracting and reporting the fluid-attenuated inversion recovery (FLAIR) signal KF on brain MRI scans (16). A traditional quantitative approach might be based on histogram analysis of normal brain FLAIR SIs. After appropriate preprocessing steps, a histogram of SI of brain voxels from MRI scans of a normal population can be described by Gaussian distribution with preliminary ± 2 SD normal/abnormal thresholds, as for conventional clinical pathology tests (17). Those voxels in the >2 SD tail of the distribution can subsequently be classified as Increased SI; the voxels <2 SD as Decreased SI; with the remainder of the voxels labeled as Normal. By this process, each voxel has a number directly reflecting the measurement of SI and a categorical label based on its SI relative to the mean of the distribution of all voxel SIs. While useful for many image analysis tasks, this analytical approach has weaknesses in the face of noise, which is present on every image. Differentiating signal from noise is difficult for these linear models.

The alternative classification approach requires the labeling, or "annotating," of brain voxels as *Increased, Normal*, or *Decreased* FLAIR SI in a training case set. This labeling is often performed by human experts and is tedious. This training set is then used to build a digital KF pattern of normal and abnormal FLAIR SI. This task can be performed by a convolutional neural network of the 3-D U-Net type, using "deep learning" artificial intelligence algorithms (18). After validation on a separate case set, this FLAIR "widget" can be applied to clinical cases to extract the FLAIR KF. These non-linear, neural network classifiers often handle image noise better than linear models, better separating the "chaff from the wheat." Note the fundamental difference of the two approaches. One is qualitative, based on the matching of human and computer categorical classification, while the other is quantitative, based on the statistical analysis of a distribution of signal measurements.

For most medical images, there is a single signal measured for each image type and, therefore, a separate computational algorithm, or "widget," is needed for each image type or modality. For a CT scan of the brain, only a single signal widget is needed to measure or classify radiodensity. For a multi-modality MRI examination, not only are signal specific pulse sequences required, but signal specific analytic widgets are necessary for FLAIR, T2, T1, diffusion-weighted imaging, susceptibility, etc. Regardless, rather than a radiologist's often ambiguous free-text report, the computer derived signal KFs are discrete and easily entered into a KF table.

It should be noted that KFs reported in this fashion are associated with only one lesion, and this is a significant limitation of this simplistic approach. If there are multiple similar appearing lesions from the same disease (metastasis), this limitation is significantly mitigated by the additional spatial KF of multiplicity. However, if there are multiple lesions from different diseases, separate analysis for each disease must be performed and reported. This is a difficult task even for humans, and is, at present, beyond computational techniques.

As with human observers, specific objects within images, such as a tumor, are detected and partially characterized on the basis of their abnormal SI. Lesions that have no abnormal signal are rare and difficult to identify. Once a computer has identified an object by its signal characteristics, whether by classification or numeric methods, the spatial features of the object must also be extracted. This requires spatial operators that combine voxels of related signal characteristics into individual objects that other algorithms must then count, measure, spatially describe, and anatomically localize. These KFs can be entered into the spatial components of a KF table.

As with radiologists, organ-based analysis is advantageous and easily performed by computers. Requirements for the evaluation of whole organ spatial patterns are "normal" anatomic atlases and computer algorithms for identifying specific organs and comparing their spatial properties to those of normal atlas templates. Remarkable progress has been made over the past 10 years in the development and use of digital, three-dimensional anatomic templates (19). Typically, tissue segmentation algorithms are applied, oftentimes relying on machine learning models. Atlas-based deformable registration methods then apply spatial transformations to the image data to bring anatomically corresponding regions into spatial co-registration with the normal atlas. There are numerous sophisticated software

programs that perform these functions for evaluating the spatial properties of an organ or lesion (20, 21). The output of these algorithms are the same spatial KFs reported by radiologists, including the number, size, and shape of organs and lesions and their anatomic locations.

The computer, by extracting brain image KFs and reporting them numerically or categorically, can generate a highly structured Findings section of a radiology report that is directly comparable to that generated by a radiologist. The computer's extracted, discrete KFs can also be entered into a computational inference engine, of which there are many. One could use simple, naïve Bayesian networks, which structurally have an independent node for every disease with conditional nodes for each KF (22,23). These tools include look-up tables with rows listing all possible diagnoses, columns for all extracted KFs, and cells containing the probabilities of KF states conditioned on each covered disease. Given a set of KFs of a clinical examination, a Bayesian network calculates the probability of each disease and ranks them into the differential diagnoses that can be incorporated into the "Impression" section of the computer report. This is a form of computational pattern recognition resulting from best matches of particular KF patterns with a specific diagnosis.

The preceding approach to computer image analysis closely resembles that of the cognitive, slow thinking, human. While the process is relatively transparent and comprehensible, it can be computationally challenging. But as with humans, there are alternative, faster thinking, heuristic computational methods, most commonly based on neural networks, that are a revolution in digital image analysis. The algorithms are usually non-linear classifiers that are designed to output a single diagnosis, and nothing else. These programs are trained on hundreds or thousands of carefully "annotated" cases, with and without the specified disease. No intermediate states or information are used or generated. In other words, there are no KFs that might inform the basis of a diagnosis, nor is there quantitative output to provide more specific information about the disease or to guide clinical management. These "black box" systems resemble human professionals thinking fast, but with little obvious insight. However, an experienced radiologist incorporates thousands of these heuristic black boxes into his/her decision-making, many of which incorporate nonimage data from the electronic medical record, local practice mores, the community, and environment.

For a computer algorithm to mimic the radiologist in daily practice, it too must incorporate thousands of widgets and vast quantities of diverse data. Such a task may not be impossible, but it does not seem eminent. Furthermore, a radiologist can, when necessary, switch from heuristics to the deliberative mode and "open" the box to explain why they made a particular diagnosis. This often involves the explication of associated KFs (mass effect) that may simultaneously be important for clinical management (decompression).

## CONCLUSION

A computer using contemporary computational tools functionally resembling human behavior could, in theory, read in image data as it comes from the scanner, extract KFs, find matching diagnoses, and integrate both into a standardized radiology report. The computer could populate the report with additional quantitative data, including organ/lesion

volumetrics and statistical probabilities for the differential diagnosis. We predict that within 10 years this conjecture will be reality in daily radiology practice, with the computer operating at the level of subspecialty fellows. Both will require attending oversight. A combination of slow and fast thinking is important for radiologists and computers.

## Abbreviations:

| | |
|---|---|
| **KF** | key feature |
| **SI** | signal intensity |
| **CT** | computed tomography |
| **MRI** | magnetic resonance imaging |
| **FLAIR** | fluid-attenuated inversion recovery |

## REFERENCES


1. Bryan RN. Introduction to the science of medical imaging. Cambridge University Press, 2010.
2. Grenander U, Miller MI. Pattern theory: from representation to inference. Oxford; New York: Oxford University Press, 2007.
3. Wandell BA, Dumoulin SO, Brewer AA. Visual field maps in human cortex. Neuron 2007; 56(2):366–383. doi:10.1016/j.neuron.2007.10.012. [PubMed: 17964252]
4. Wandell BA. Foundations of vision. Sinauer Associates, 1995.
5. Zeki S A vision of the brain. Blackwell Scientific Publications, 1993.
6. Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev 1956; 63(2):81–97. doi:10.1037/h0043158. [PubMed: 13310704]
7. Langlotz CP. The radiology report: a guide to thoughtful communication for radiologists and other medical professionals. 1st ed CreateSpace Independent Publishing Platform, 2015.
8. Reeder MM. Gamuts in radiology - 4.0 web edition. Available at: http://gamuts.isradiology.org/Gamuts.htm. Accessed July. 20 2019
9. STATdx. Available at: https://www.statdx.com/. Accessed July 20 2019.
10. Thinking Kahneman D., fast and slow. 1st ed New York: Farrar, Straus and Giroux, 2011.
11. Carmody DP, Nodine CF, Kundel HL. An analysis of perceptual and cognitive factors in radiographic interpretation. Perception 1980; 9(3):339–344. doi:10.1068/p090339. [PubMed: 7454514]
12. Gladwell M Outliers: the story of success. 1st ed New York: Little, Brown and Co, 2008.
13. Busby LP, Courtier JL, Glastonbury CM. Bias in radiology: the how and why of misses and misinterpretations. Radiographics 2018; 38(1):236–247. doi:10.1148/rg.2018170107. [PubMed: 29194009]
14. Epstein CL. Introduction to the mathematics of medical imaging. Upper Saddle River, N.J: Pearson Education/Prentice Hall, 2003.
15. Larson DB, Chen MC, Lungren MP, et al. Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. Radiology 2018; 287(1):313–322. doi:10.1148/radiol.2017170236. [PubMed: 29095675]
16. Jack CR, O'Brien PC, Rettman DW, et al. FLAIR histogram segmentation for measurement of leukoaraiosis volume. J Magn Reson Imaging 2001; 14(6). 668–376. [PubMed: 11747022]
17. Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford University Press, 2004.

18. Duong MT. Convolutional neural network for automated FLAIR lesion segmentation on clinical brain MR imaging. Am J Neuroradiol 2019. (in press).

19. Doshi J, Erus G, Ou Y, et al. MUSE: MUlti-atlas region segmentation utilizing ensembles of registration algorithms and parameters, and locally optimal atlas selection. Neuroimage 2016; 127:186–395. doi:10.1016/j.neuroimage.2015.11.073.

20. Shen D, Davatzikos C. HAMMER: hierarchical attribute matching mechanism for elastic registration. IEEE Trans Med Imaging 2002; 21(11):1421–1439. doi:10.1109/TMI.2002.803111. [PubMed: 12575879]

21. Davatzikos C, Genc A, Xu D, et al. Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. Neuroimage 2001; 14(6):13611369. doi:10.1006/nimg.2001.0937.

22. Chen R, Wang S, Poptani H, Herskovits EH, et al. A Bayesian diagnostic system to differentiate glioblastomas from solitary brain metastases. Neuroradiol J 2013; 26(2):175–183. doi:10.1177/197140091302600207. [PubMed: 23859240]

23. Herskovits EH, Cooper GF. Algorithms for Bayesian belief-network pre-computation. Methods Inf Med 1991; 30(2):81–89. [PubMed: 1857253]