

# Imaging-based observational databases for clinical problem solving: the role of informatics

Alex A T Bui, William Hsu, Corey Arnold, Suzie El-Saden, Denise R Aberle, Ricky K Taira

Medical Imaging Informatics (MII) Group, Department of Radiological Sciences, UCLA David Geffen School of Medicine, Los Angeles, California, USA

## Correspondence to

Dr Alex A T Bui, Medical Imaging Informatics (MII) Group, Department of Radiological Sciences, UCLA Medical Imaging Informatics, 924 Westwood Blvd, Suite 420, Los Angeles, CA 90024, USA; buia@mii.ucla.edu

Received 14 September 2012

Revised 3 May 2013

Accepted 3 June 2013

Published Online First

17 June 2013

## ABSTRACT

Imaging has become a prevalent tool in the diagnosis and treatment of many diseases, providing a unique in vivo, multi-scale view of anatomic and physiologic processes. With the increased use of imaging and its progressive technical advances, the role of imaging informatics is now evolving—from one of managing images, to one of integrating the full scope of clinical information needed to contextualize and link observations across phenotypic and genotypic scales. Several challenges exist for imaging informatics, including the need for methods to transform clinical imaging studies and associated data into structured information that can be organized and analyzed. We examine some of these challenges in establishing imaging-based observational databases that can support the creation of comprehensive disease models. The development of these databases and ensuing models can aid in medical decision making and knowledge discovery and ultimately, transform the use of imaging to support individually-tailored patient care.

## INTRODUCTION

A key goal of imaging informatics research is in supporting medical knowledge discovery from the wealth of routinely acquired clinical imaging and associated data, resulting in systems that can better inform, individualize, and optimize patient care. We briefly present one perspective, largely from the viewpoint of radiology, toward this objective through the creation of imaging-based observational databases that can enable disease modeling and drive applications. Indeed, a longstanding pursuit of biomedical informatics has been the automated creation of observational databases.<sup>1–2</sup> Efforts like STRIDE and i2b2 have helped establish clinical repositories that have been used in a number of applications, including cohort identification, detection of drug interactions, and population-level analyses.<sup>3–4</sup> Analogous applications are envisioned atop imaging-based observational databases, such as similarity search (for cohort identification), knowledge discovery, and population-level studies. But just as clinical data warehouses have encountered issues with data standardization and the need for common representations (along with appropriate mapping methods), an imaging-based observational database faces similar issues. Imaging data and its acquisition pose a unique set of requirements and challenges, mainly around the need for rich interpretive context (potentially from other data sources) to properly understand findings. While a number of infrastructure tools exist for storing and sharing imaging datasets (eg, Extensible Neuroimaging Archive Toolkit

(XNAT),<sup>5</sup> the National Cancer Imaging Archive (NCIA), Transfer of Images and Data (TRIAD)), these tools do not directly address challenges that affect the formal, automated analysis of available images. Markedly, the incomplete recording of the ‘context’ surrounding clinical observations impedes the application of more advanced computational techniques and the secondary use of these collected data. An ongoing issue is the discrepancy between our growing capacity to gather patient data without a similar increase in our ability to analyze, understand, and act rationally on this information.<sup>6</sup> These problems can be overcome by imaging informatics developments, transforming the data into information that can enable scientific-quality models. In turn, these integrative imaging-based models can be instantiated to empower decision-making tailored for each individual. This idea motivates figure 1 and the sections below, which extend a (simplified) imaging workflow to illustrate how novel informatics-based methods can be used to better characterize patient data and construct an observational database.

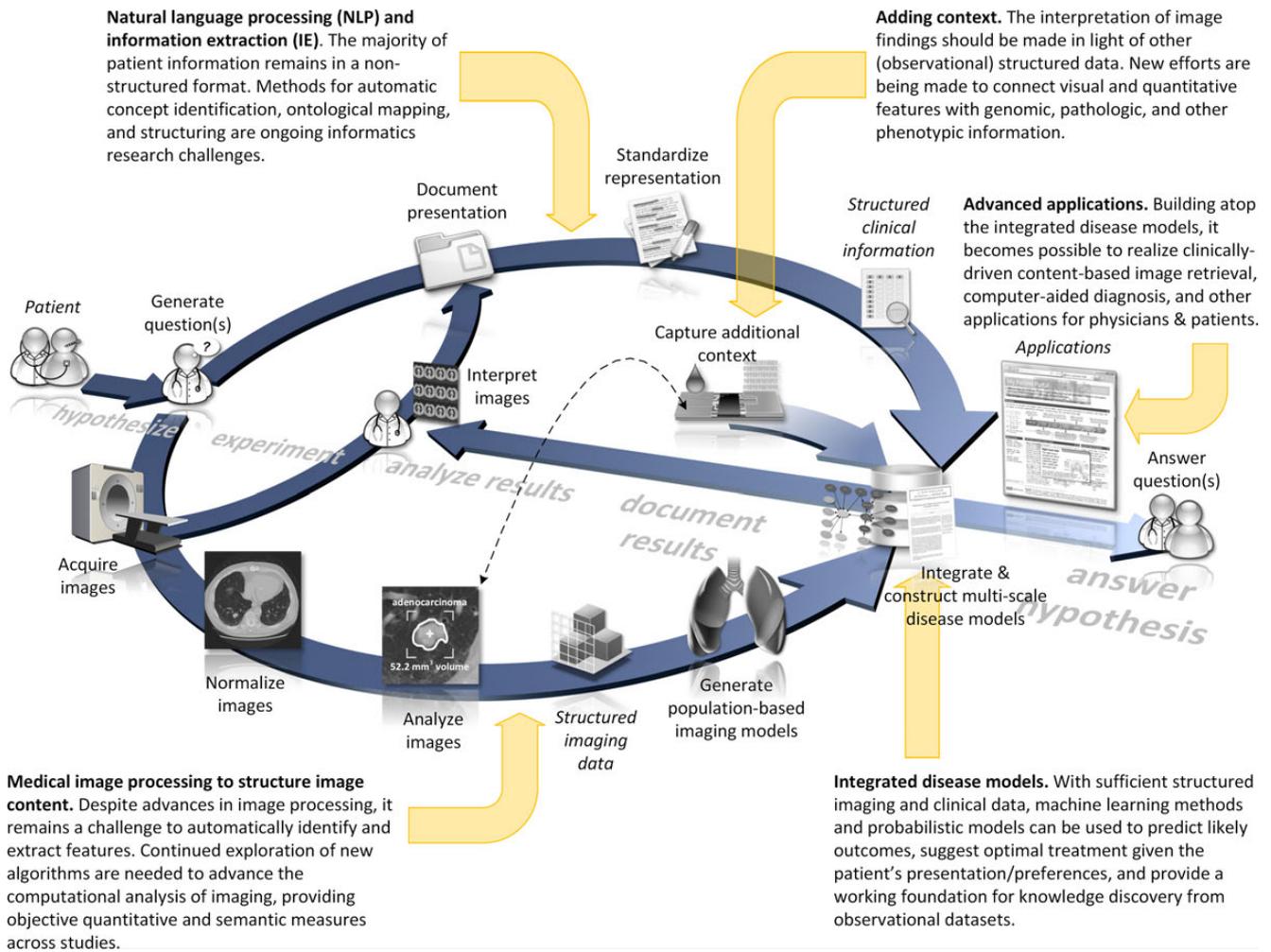
## CHALLENGES IN BUILDING IMAGING-BASED OBSERVATIONAL DATABASES

### Improving image/study characterization

Traditionally, imaging interpretation has been the purview of the radiologist, providing an understanding based on the identification of visual patterns in relation to the patient’s presentation and clinical history. In an abstract manner, the radiologist mentally constructs an anatomic/physiologic representation of the individual from a projectional or cross-sectional study, implicitly considering the quality of the study and prior imaging findings; and comparing radiographic features against a working knowledge of similar cases.<sup>7</sup> The radiologist’s undertaking is part visual search, part diagnostic reasoning.<sup>8</sup> The bottom of figure 1 illustrates steps toward structuring and making image content comparable, facilitating contextualization and automated understanding:

1. *Normalizing images*: The first step is the calibration of values obtained from study image acquisition. Such standardization is necessary so that observed changes between images (eg, in the same individual to periodically assess therapeutic response; among subjects from studies conducted at different institutions) are attributable to morphologic/physiologic change rather than acquisition differences (eg, scanner variations, environment). Normalization involves signal correction (eg, histogram equalization, T1/T2 re-estimation from magnetic resonance

**To cite:** Bui AAT, Hsu W, Arnold C, et al. *J Am Med Inform Assoc* 2013;20:1053–1058.



**Figure 1** A high-level framework for the construction of imaging-based observational databases. Central to the development of such a database is the improved contextualization of the imaging study and related patient data. The bottom half illustrates steps necessary to make imaging studies structured, including normalization, quantitative/semantic analysis, and subsequent population-based imaging constructs to facilitate comparisons. The top half of shows that information must be extracted from associated clinical (text) reports. Collectively, an integrated database can then be instantiated, from which disease models and applications can be built.

(MR) parameters, standardized uptake values for fluoro-deoxyglucose positron emission tomography (PET)), image processing (eg, denoising), and spatial regularization (eg, intra-subject registration). Importantly, as new types of imaging and protocols are introduced, methods must be in place to properly adjust observed values to accepted standards. Noteworthy efforts toward protocol and image standardization include: the workshop convened by the National Institute of Standards and Technology, examining measurement uncertainties inherent to biomarkers derived from medical images<sup>9</sup>; and the American College of Radiology's Uniform Protocols for Imaging in Clinical Trials (UPICT), which is developing a standardized template for representing and sharing imaging protocols.<sup>10</sup> Other efforts, such as the Lung Imaging Database Consortium (LIDC), the National Lung Screening Trial (NLST), and the Alzheimer's Disease Neuroimaging Initiative (ADNI)<sup>11-13</sup> have established standard protocols for specific domains and/or targeted tasks (eg, lung nodule detection). Nevertheless, these efforts have primarily focused on protocols for clinical trials or controlled data collection; standards have yet to be established that would ensure reproducibility across different institutions'

routine clinical environments. Alternatively, techniques that can automatically remap voxel-level values to calibrated references can also provide a means to normalize clinical studies across routine acquisitions.

2. *Analyzing images:* The next step entails characterizing features and identifying higher-level abstractions within images. While radiologists learn to perceive and assign the importance of a given set of image features, the development of image analysis algorithms that can handle the subtleties arising in clinical imaging remains an open area of research.<sup>14</sup> Image analysis spans a gamut, from shape and texture analyses, which derive quantitative metrics; to more sophisticated segmentation and classification schemes that attach semantic labels with regions of interest. The results of this step are objective, structured measures describing image findings. Multimodal imaging and image fusion is also considered in this step, wherein the combination of different imaging modalities provides extra feature characterization and context (eg, MR- or CT-PET).<sup>15</sup> The applications of these imaging features are limited without the ability to share and repeat such measures across different scanners and institutions. The variation introduced by (automated)

measurements (eg, by different algorithms assessing the same object) and feature extraction algorithms need to be captured to recognize potential limitations, biases, and errors; and to further comparison and reproducibility. Annotation Image Markup (AIM)<sup>16</sup> and DICOM (Digital Imaging and Communication in Medicine) Structured Reporting (DICOM-SR)<sup>17</sup> exemplify continuing efforts to improve the interoperability of radiological findings generated by various imaging toolkits and workstations.

3. *Generating population-based imaging models*: Radiologists successfully derive conclusions across dissimilar studies every day by comparing images against a mental construct they have learned over time, assessing new studies against those seen previously.<sup>18</sup> Similarly, computational models are needed to perform population-based comparisons of values generated from imaging analyses. Aggregate shape models and atlases are demonstrative, providing contemporary understandings in targeted cohorts<sup>19–21</sup>; however, additional efforts are required to provide coverage across different diseases and populations.

Critically, these techniques must be made practical in the clinical setting, which is prone to non-ideal acquisition and real-world constraints of limited scan time and variable imaging platforms. These techniques must also be computationally tractable, running in a timely manner so physicians can act on results. Underlying data standards, such as DICOM, may need further updating to fully enable the above steps, addressing the current limitations given the introduction of new imaging and data types, and increasing data sizes. Through this framework, algorithms to assess image differences can affect the transition from subjective interpretations to objective, reproducible characterization of findings and patient state.<sup>22</sup>

### Placing findings and the patient into context

Image structuring must be paralleled by methods for related data that add insight into a patient's condition, as depicted by the top half of figure 1. The preponderance of non-imaging observations remains in free-text reports. For example, having information about active medications that a patient is taking and other past interventions (eg, surgical resection for a tumor) is imperative in assessing treatment response as quantified from follow-up imaging studies. Often, the interpretation of an imaging finding is influenced by documented patient history,<sup>7 18</sup> with information coming not only from prior radiology reports but also from a range of clinical notes including admission/discharge summaries, surgical reports, specialty consults (eg, oncology, neurology, urology, etc.), and pathology reports. While structured reporting<sup>23–26</sup> and the use of controlled vocabularies (eg, Unified Medical Language System, RadLex) have been pursued, including to enable secondary usage of clinical documents, the majority of report generation is still designed to accommodate the physician and the natural expression of language. At its core, the extraction and linking of patient context from these reports requires identification and contextualization of semantic concepts within documents. Markedly, medical natural language processing (NLP) has been applied to different medical reports to extract and structure information,<sup>27</sup> which can be used to inform image interpretation. Though well-known approaches (eg, regular expressions, rule-bases, statistical models, linguistic operations) have been applied to uncover concepts in clinical text, establishing contextual models of patients and disease will require moving beyond document-level information extraction. Continuing challenges in medical NLP for (but not limited to) radiology include co-reference resolution

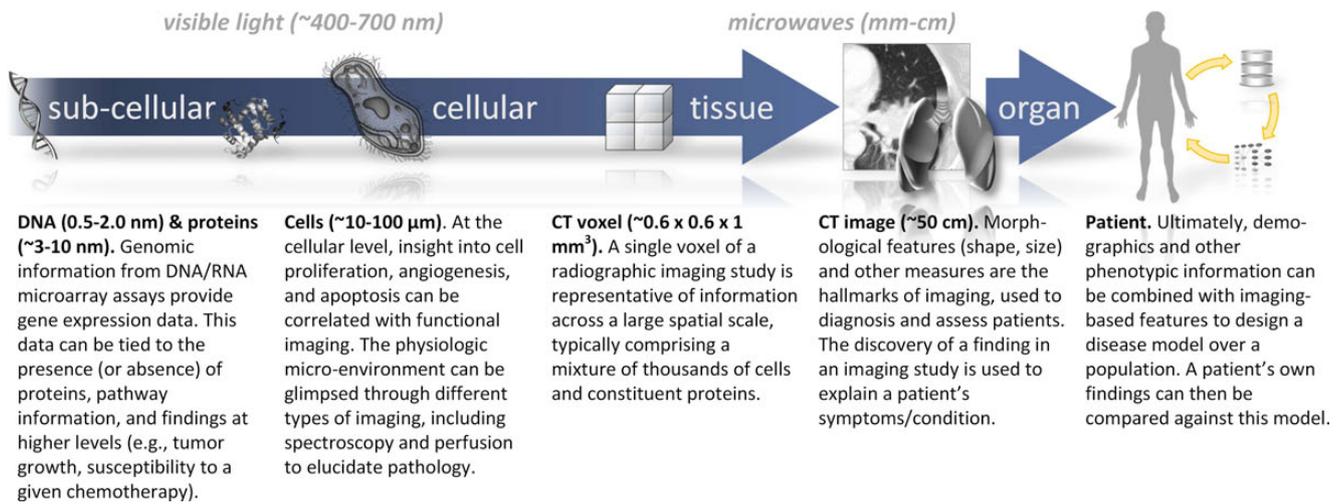
(eg, tracking image findings over time)<sup>28 29</sup>; temporal topic modeling (eg, monitoring changing patterns of semantic themes)<sup>30 31</sup>; and entailment (eg, recognizing that the mention of a resection cavity implies prior surgery).<sup>32</sup> Complementary to NLP efforts are methods for the logical representation of observations, inferences, interventions, and recommendations found in clinical text.<sup>33 34</sup> These efforts are seen in the definition of semantic relations; and situational and domain-specific ontologies. Curation of these ontologies can be assisted by NLP, data mining, and topic modeling approaches.<sup>35 36</sup>

### Bringing it together

Once image content is characterized and associated data recast into a standardized representation, it becomes feasible to combine the information into a database supporting deeper integration of a patient's symptoms, findings, treatment, and long-term outcomes. If we consider the physical resolution of one voxel (figure 2) from a radiographic imaging study, it typically consists of an assortment of thousands of cells, which in turn are the product of expressed genetic material. It is at this point where imaging becomes a gateway to multi-scale models of disease, relating phenotypic presentation with molecular, cellular, and other physiologic findings.

### Linking across biological scales: imaging biomarkers

In the past decade there has been rising interest in the development of imaging-based biomarkers.<sup>37</sup> As suggested in figure 1 by the dashed arrow, there is reciprocal value in determining the relationships between image-derived features and disease indicators at the cellular and genetic levels. Functional imaging (eg, PET, dynamic contrast-enhanced MR, MR spectroscopy, perfusion) is being explored across biological scales and diseases. Features from such studies have been correlated with properties at the cellular (eg, proliferation, angiogenesis) and genetic (eg, gene expression, single nucleotide polymorphism) levels.<sup>38–40</sup> Quantitative imaging biomarkers can be incorporated into parametric maps summarizing the spatial differences related to gene expression, local environment, and cellular function.<sup>41</sup> These maps yield knowledge about disease pathophysiology and treatment response in neurological disorders and oncology.<sup>42–45</sup> Despite these important insights, the clinical application of imaging biomarkers has been limited. Given the number of comparisons performed and the potential for erroneous associations, new imaging biomarkers often meet with skepticism.<sup>46</sup> Consortia such as the NCIA, Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA), Osteoarthritis Initiative (OAI), and Quantitative Imaging Network (QIN) have amassed large numbers of imaging studies to validate and reproduce findings.<sup>10 47–50</sup> Concerted efforts such as The Cancer Genome Atlas (TCGA) and projects with biospecimen repositories (eg, NLST) that acquire genomic and related information (clinical, imaging) are beginning to generate important insights.<sup>51</sup> However, the richness of associated clinical history varies greatly (eg, a single outcome is reported in the dataset, or only a single time point over the course of disease; limited information captured by a study case report form), thereby decreasing the opportunity to fully mine results. Indeed, variability arising from these datasets' origins (eg, routine imaging exams versus clinical trials) contributes to the difficulty of transforming this data into clinically useful applications. This challenge underscores the need to improve on current methods for imaging and study characterization and provide a means for integrating information across data sources and biological scales.

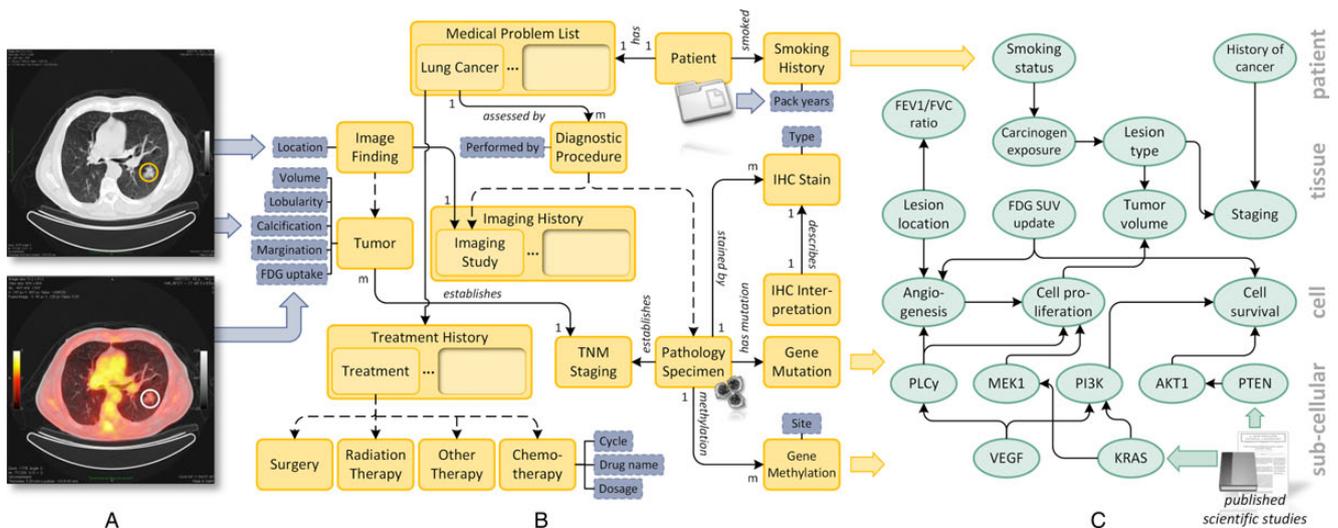


**Figure 2** Spatially, a pixel or voxel in an imaging study represents the combination of several different structural and physiologic elements. For instance, a single image in an axial CT study has a field of view of ~50 cm; and a single voxel within this image has typical dimensions of 0.6×0.6×1 mm<sup>3</sup>, in which thousands of cells may be represented. New types of imaging are elucidating sub-cellular and cellular processes. As such, imaging provides a multi-scale perspective on a disease.

Constructing a disease model

Given a database of structured imaging and patient data, models can be assembled to organize and explain observations (figure 3), and serve as the basis for prognostic tools. Clinical and imaging data sources can be combined to create models that elucidate the underlying mechanisms of a disease: the generation of quantitative maps from imaging modalities (eg, PET, diffusion weighted imaging) provide estimates of biological processes like cell proliferation.<sup>38 52 53</sup> Information from these quantitative features may further be incorporated into a model that predicts overall survival or disease progression using logistic regression or machine learning methods (eg, neural networks, random forests, support

vector machines).<sup>54</sup> For instance, one such modeling formalism is the Bayesian belief network (BBN),<sup>55–57</sup> which is a directed acyclic graph of nodes representing random variables and edges indicating conditional dependence (figure 3C). By way of illustration, a BBN for predicting lung cancer based on screening data might comprise variables capturing image features (eg, nodule size, texture features), demographics and social history (eg, smoking history, family history of cancer), genetics (eg, markers associated with lung cancer), and related conditions (eg, other pulmonary diseases that increase risk for lung cancer); links between these variables indicate influence or causal relations. From the integrated database, it becomes possible to: define the



**Figure 3** Integrating imaging and clinical information to construct a disease model. (A) Features are computed from normalized imaging studies. In this case, a lung cancer is identified on a CT imaging study (top image), and a corresponding positron emission tomography study (bottom image) is performed to assess the level of tumor activity. (B) An integrated data model creates the basis for combining image data with additional information from the patient’s medical record, thereby providing context to the imaging study interpretation. For instance, information on the patient’s smoking history, results of thoracic tumor biopsies, and current (and past) treatments lend insight into an individual’s status. (C) A multi-scale disease model can be populated from this imaging-based observational database to establish a population-based perspective on the condition (a partial model is shown). Conversely, if a disease model has been vetted for a given population, then the model can be used as a prognostic tool for a given individual, offering tailored recommendations. Knowledge from published scientific studies and other resources can also be embedded within the disease model.

scope of variables of (clinical) interest that are pertinent to a given disease; infer a network topology (or establish one given expert knowledge) for the BBN; and calculate the conditional probabilities needed per network node. The resultant BBN embodies the joint distribution across the set of observations for a population, providing a representation that can be queried to facilitate medical decision-making tasks, including optimizing a given individual's care. To date, however, few models have been widely adopted; notably, difficulties lie in the generalized application of these models. Although such models may leverage a large population, the findings may still be site- and/or population-specific. Also, we increasingly see studies that span time, particularly when long-term outcomes are involved; as such, the nature of observations and interventions can change, especially with introduction of new interventional or therapeutic means (eg, approval for a new surgical device or chemotherapy). Hence, these integrative models need to begin reflecting time to enable a complete understanding of patients and the environs in which decisions are made.

### Connecting biomedical research

Observational databases provide a foundation on which to construct disease models, but remain potentially limited to the characteristics of populations and standards of care of the institution(s) from which the data originate. In contrast, the scientific literature affords insight into variables and pathways across institutional cultures. But as noted by a recent Institute of Medicine report, overwhelming amounts of medical information are continuously published via clinical trials, journals, and other vetted resources; it is the role of informatics to provide real-time access to this knowledge.<sup>58</sup> Such efforts pose their own challenges of standardization, as the circumstances, external validity, and description of reported populations vary.<sup>59</sup> As with imaging and clinical documentation, the contents of biomedical literature must be mapped into a formal representation.<sup>60</sup> Having structured this information, a disease model that integrates this evidence in a principled manner can be realized to help clinicians.<sup>61</sup>

## DISCUSSION

This perspective has touched on how different parts of imaging and informatics pursuits can be brought together to create an imaging-based observational database from clinical data. The described issues shape some of the enduring and future research challenges for imaging informatics. Admittedly, with the breadth of research and development in the many areas comprising imaging informatics, we have only been able to touch on some of the issues and provide a few examples. And while radiologic imaging has been our focus, the cited challenges are germane to other domains employing imaging, like pathology.

Though elements of past and current imaging informatics efforts separately address some issues in focused domains, it is important that a synergistic, common framework come about—much like with the origins of the Insight Toolkit (ITK),<sup>62</sup> which pulled together multiple groups under the auspices of one coordinating entity to create a shared vision and set of open tools—to support dissemination and a broader user community. While such efforts have occurred to some extent in the above cited works, a more generalized approach is now warranted. At the same time, while data standards such as DICOM have attempted to keep pace with data representational issues, it is often after the fact; a forward-looking revision of this and other related standards should take into account the information requisite to support image and image feature normalization,

consolidating lessons learned, current approaches, and best practices together. The imaging informatics community plus vendors need to quickly implement this change. Such a framework would allow a continuous stream of imaging cases to be used for characterization of normal and diseased populations. Groups such as the Society for Imaging Informatics in Medicine (SIIM) and the American College of Radiology Imaging Network (ACRIN), which have long advocated for core imaging informatics standards and projects, can play a central role in such an endeavor along with federal funding agencies like the National Institutes of Health (NIH). Many of the issues highlighted in transforming data into normalized forms are to alleviate barriers to comparison and secondary data use. Currently, informatics-based techniques are applied after the fact in an attempt to impose structure and restore context: if full study characterization and structuring occurs upfront, many issues could be resolved and the construction of disease models fostered. We are only at the technical cusp of realizing the full promise of imaging repositories and the electronic medical record, with databases that can support evidence-based medicine and personalize treatment through a range of applications including more sophisticated content-based image retrieval; computer-aided detection/diagnosis; and knowledge discovery via imaging to differentiate disease subtypes.<sup>63–66</sup>

Imaging informatics is evolving: while aspects have focused on its exclusive role within radiology to improve image processing and workflow, the field and its research has an opportunity to become the nucleus for organizing, analyzing, and communicating the breadth of information involved in diagnosis and treatment. And ultimately, the potential for imaging informatics is vast: as a nexus for information across clinical domains, it is one of the few frameworks that can provide a multi-scale, in vivo perspective on a given patient's condition and outcome.<sup>67–69</sup> Essential to this transformation is a holistic view of imaging informatics, appreciating its intersection with developments from other disciplines. In turn, this change will help reinvent the role of radiology as a collaborative activity that guides and impacts patient healthcare.<sup>70</sup>

**Acknowledgements** The authors would like to thank Dr Hooshang Kangarloo for his insights and vision, which have guided the UCLA Medical Imaging Informatics Group over many years.

**Contributors** AB formed the outline for this perspective, and was responsible for the writing and integration of material from the other authors. WH, CA, and RT provided additional writing and materials for this article. SE and DA provided clinical insight, comments, and further editing of the manuscript.

**Funding** Aspects of the authors' work have been funded by National Institutes of Health (NIH) grants R01 CA157553, R01 NS076534, R01 EB000362, R01 LM011333, and T15 LM007536.

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Hlatky MA, Lee KL, Harrell FE, *et al.* Tying clinical research to patient care by use of an observational database. *Stat Med* 1984;3:375–84.
- Hlatky MA. Using databases to evaluate therapy. *Stat Med* 1991;10:647–52.
- Lowe HJ, Ferris TA, Hernandez PM, *et al.* STRIDE—an integrated standards-based translational research informatics platform. *Proceedings AMIA Fall Symposium* 2009:391–5.
- Gainer V, Hackett K, Mendis M, *et al.* Using the i2b2 hive for clinical discovery: an example. *Proceedings AMIA Fall Symposium* 2007:959.
- Marcus DS, Olsen TR, Ramaratnam M, *et al.* The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. *Neuroinformatics* 2007;5:11–34.
- Andriole KP, Wolfe JM, Khorasani R, *et al.* Optimizing analysis, visualization, and navigation of large image data sets: one 5000-section CT scan can ruin your whole day. *Radiology* 2011;259:346–62.

- 7 Tagare HD, Jaffe CC, Duncan J. Medical image databases: a content-based retrieval approach. *J Am Med Inform Assoc* 1997;4:184–98.
- 8 Donovan T, Manning DJ. The radiology task: Bayesian theory and perception. *Br J Radiol* 2007;80:389–91.
- 9 Clarke LP, Sriram RD, Schilling LB. Imaging as a biomarker: standards for change measurements in therapy workshop summary. *Acad Radiol* 2008;15:501–30.
- 10 Buckler AJ, Bresolin L, Dunnick NR, et al. A collaborative enterprise for multi-stakeholder participation in the advancement of quantitative imaging. *Radiology* 2011;258:906–14.
- 11 McNitt-Gray MF, Armato SG III, Meyer CR, et al. The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation. *Acad Radiol* 2007;14:1464–74.
- 12 National Lung Screening Trial ResearchAberle T, Berg DR, Mueller SG, Weiner MW, Thal LJ, et al. The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clin N Am* 2005;15:869–77, xi–xii.
- 14 Deserno T, Aach T, Amunts K, et al. Advances in medical image processing. *Comput Sci Res Dev* 2011;26:1–3.
- 15 Fei B, Yang X, Nye JA, et al. MRPET quantification tools: registration, segmentation, classification, and MR-based attenuation correction. *Med Phys* 2012;39:6443–54.
- 16 Channin DS, Mongkolwat P, Kleper V, et al. The caBIG™ annotation and image markup project. *J Digit Imaging* 2010;23:217–25.
- 17 Clunie DA. DICOM structured reporting and cancer clinical trials results. *Cancer Inform* 2007;4:33.
- 18 Nyce JM, Steele JS, Gunderman RB. Bridging the knowledge divide in radiology education. *Radiology* 2006;239:629–31.
- 19 Cootes TF, Edwards GJ, Taylor CJ. Active appearance models. *IEEE Trans Pattern Anal Mach Intell* 2001;23:681–5.
- 20 Pizer SM, Fletcher PT, Joshi S, et al. Deformable M-reps for 3D medical image segmentation. *Int J Comput Vis* 2003;55:85–106.
- 21 Toga AW, Thompson PM. What is where and why it is important. *Neuroimage* 2007;37:1045–9.
- 22 Boone JM. Radiological interpretation 2020: toward quantitative image assessment. *Med Phys* 2007;34:4173–9.
- 23 Bell DS, Greenes RA, Doubilet P. Form-based clinical input from a structured vocabulary: initial application in ultrasound reporting. *Proceedings AMIA Fall Symposium* 1992:789.
- 24 Liberman L, Menell JH. Breast imaging reporting and data system (BI-RADS). *Radiol Clin North Am* 2002;40:409–30.
- 25 Hussein R, Engelmann U, Schroeter A, et al. DICOM structured reporting: Part 1. Overview and characteristics. *Radiographics* 2004;24:891–6.
- 26 Reiner B. The challenges, opportunities, and imperative of structured reporting in medical imaging. *J Digital Imaging* 2009;22:562–8.
- 27 Meystre SM, Savova GK, Kipper-Schuler KC, et al. Extracting information from textual documents in the electronic health record: a review of recent research. *Yearb Med Inform* 2008;1:128–44.
- 28 Son RY, Taira RK, Kangaroo H. Inter-document coreference resolution of abnormal findings in radiology documents. *Stud Health Technol Inform* 2004; 107:1388–92.
- 29 Zheng J, Chapman WW, Crowley RS, et al. Coreference resolution: a review of general methodologies and applications in the clinical domain. *J Biomed Inform* 2011;44:1113–22.
- 30 Arnold C, Speier W. A topic model of clinical reports. *Proceedings ACM SIGIR* 2012:1031–2.
- 31 Wang X, McCallum A. Topics over time: a non-Markov continuous-time model of topical trends. *Proceedings ACM SIGKDD* 2006:424–33.
- 32 Bos J, Markert K. Recognising textual entailment with logical inference. *Proceedings Association Comparative Linguistics* 2005:628–35.
- 33 Ramos DR, ten Teije A, Miksch S, et al. *Knowledge representation for health-care*. Berlin: Springer-Verlag, 2011.
- 34 Coden A, Savova G, Sominsky I, et al. Automatically extracting cancer disease characteristics from pathology reports into a disease knowledge representation model. *J Biomed Inform* 2009;42:937–49.
- 35 Krallinger M, Valencia A, Hirschman L. Linking genes to literature: text mining, information extraction, and retrieval applications for biology. *Genome Biol* 2008;9 (Suppl 2):S8.
- 36 Arnold CW, El-Saden S, Bui AAT, et al. Clinical case-based retrieval using latent topic analysis. *Proceedings AMIA Fall Symposium* 2010:26–30.
- 37 Smith JJ, Sorensen AG, Thrall JH. Biomarkers in imaging: realizing radiology's future. *Radiology* 2003;227:633–8.
- 38 Padhani AR, Miles KA. Multiparametric imaging of tumor response to therapy. *Radiology* 2010;256:348–64.
- 39 Rutman AM, Kuo MD. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. *Eur J Radiol* 2009;70:232–41.
- 40 Thompson PM, Martin NG, Wright MJ. Imaging genomics. *Curr Opin Neurol* 2010;23:368–73.
- 41 Moffat BA, Chenevert TL, Meyer CR, et al. The functional diffusion map: an imaging biomarker for the early prediction of cancer treatment outcome. *Neoplasia* 2006;8:259–67.
- 42 Padhani AR. Diffusion magnetic resonance imaging in cancer patient management. *Semin Radiat Oncol* 2011;21:119–40.
- 43 Cyran CC, von Einem JC, Paprotka PM, et al. Dynamic contrast-enhanced computed tomography imaging biomarkers correlated with immunohistochemistry for monitoring the effects of sorafenib on experimental prostate carcinomas. *Invest Radiol* 2012;47:49–57.
- 44 Nishino M, Jackman DM, Hatabu H, et al. Imaging of lung cancer in the era of molecular medicine. *Acad Radiol* 2011;18:424–36.
- 45 Phan JH, Quo CF, Wang MD. Functional genomics and proteomics in the clinical neurosciences: data mining and bioinformatics. *Prog Brain Res* 2006;158:83–108.
- 46 Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, et al. Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. *Radiology* 2008;248:625–31.
- 47 Clarke LP, Croft BS, Nordstrom R, et al. Quantitative imaging for evaluation of response to cancer therapy. *Transl Oncol* 2009;2:195–7.
- 48 Jack CR Jr, Bernstein MA, Borowski BJ, et al. Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. *Alzheimers Dement* 2010;6:212–20.
- 49 Madhavan S, Zenklusen JC, Kotliarov Y, et al. Rembrandt: helping personalized medicine become a reality through integrative translational research. *Mol Cancer Res* 2009;7:157–67.
- 50 Van Horn JD, Toga AW. Is it time to re-prioritize neuroimaging databases and digital repositories? *Neuroimage* 2009;47:1720–34.
- 51 Gutman DA, Cooper LA, Hwang SN, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set. *Radiology* 2013;267:569–96.
- 52 Ellingson BM, Cloughesy TF, Lai A, et al. Quantitative probabilistic functional diffusion mapping in newly diagnosed glioblastoma treated with radiochemotherapy. *Neuro-oncology* 2013;15:382–90.
- 53 Principi M, Italiani M, Guiducci A, et al. Perfusion MRI in the evaluation of the relationship between tumour growth, necrosis and angiogenesis in glioblastomas and grade 1 meningiomas. *Neuroradiology* 2003;45:205–11.
- 54 Adams ST, Leveson SH. Clinical prediction rules. *BMJ* 2012;344:d8312.
- 55 Pearl J. *Causality: models, reasoning, and inference*. New York: Cambridge University Press, 2000.
- 56 Lucas PJ, van der Gaag LC, Abu-Hanna A. Bayesian networks in biomedicine and healthcare. *Artif Intell Med* 2004;30:201–14.
- 57 Donkers J, Tuyls K. Belief networks for bioinformatics. *Compu Intell Bioinform* 2008:75–111.
- 58 Smith M, Saunders R, Stuckhardt L, et al. *Best care at lower cost: the path to continuously learning health care in America*. Washington, DC: The National Academies Press, 2012.
- 59 Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005;365:1159–62.
- 60 Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med* 2010;152:726–32.
- 61 Haynes RB. Of studies, syntheses, summaries, and systems: the '5S' evolution of information services for evidence-based healthcare decisions. *Evid Based Med* 2006;11:162–4.
- 62 Yoo TS, Ackerman MJ, Lorensen WE, et al. Engineering and algorithm design for an image processing API: a technical report on ITK—The Insight Toolkit. *Stud Health Technol Inform* 2002;85:586–92.
- 63 Zhou XS, Zillner S, Moeller M, et al. Semantics and CBIR: a medical imaging perspective. *Proceedings 2008 ACM International Conference Content-based Image and Video Retrieval* 2008:571–80.
- 64 Akgul CB, Rubin DL, Napel S, et al. Content-based image retrieval in radiology: current status and future directions. *J Digit Imaging* 2011;24:208–22.
- 65 Welter P, Riesmeier J, Fischer B, et al. Bridging the integration gap between imaging and information systems: a uniform data concept for content-based image retrieval in computer-aided diagnosis. *J Am Med Inform Assoc* 2011;18:506–10.
- 66 Belden CJ, Valdes PA, Ran C, et al. Genetics of glioblastoma: a window into its imaging and histopathologic variability. *Radiographics* 2011;31:1717–40.
- 67 Depeursinge A, Muller H. Sensors, medical images and signal processing: the growth of clinicians' expectations. *Yearb Med Inform* 2011;6:92–5.
- 68 Reiner BI, Siegel EL. The clinical imperative of medical imaging informatics. *J Digital Imaging* 2009;22:345–7.
- 69 Mollura DJ, Carrino JA, Matuszak DL, et al. Bridging radiology and public health: the emerging field of radiologic public health informatics. *J Am Coll Radiol* 2008;5:174–81.
- 70 Maynard CD. Radiologists: physicians or expert image interpreters? *Radiology* 2008;248:333–6.