

# Quantitative Imaging Biomarkers: The Application of Advanced Image Processing and Analysis to Clinical and Preclinical Decision Making

Jeffrey William Prescott

Published online: 14 March 2012  
© Society for Imaging Informatics in Medicine 2012

**Abstract** The importance of medical imaging for clinical decision making has been steadily increasing over the last four decades. Recently, there has also been an emphasis on medical imaging for preclinical decision making, i.e., for use in pharmaceutical and medical device development. There is also a drive towards quantification of imaging findings by using quantitative imaging biomarkers, which can improve sensitivity, specificity, accuracy and reproducibility of imaged characteristics used for diagnostic and therapeutic decisions. An important component of the discovery, characterization, validation and application of quantitative imaging biomarkers is the extraction of information and meaning from images through image processing and subsequent analysis. However, many advanced image processing and analysis methods are not applied directly to questions of clinical interest, i.e., for diagnostic and therapeutic decision making, which is a consideration that should be closely linked to the development of such algorithms. This article is meant to address these concerns. First, quantitative imaging biomarkers are introduced by providing definitions and concepts. Then, potential applications of advanced image processing and analysis to areas of quantitative imaging biomarker research are described; specifically, research into osteoarthritis (OA), Alzheimer's disease (AD) and cancer is presented. Then, challenges in quantitative imaging biomarker research are discussed. Finally, a conceptual framework for integrating clinical and preclinical considerations into the development of quantitative imaging biomarkers and their computer-assisted methods of extraction is presented.

**Keywords** Image analysis · Image processing · Decision making · Quantitative imaging biomarkers

## Introduction

The overwhelming success of medical science and industry in the 20th century in delivering innovative discoveries to patients has recently given way to a relative stagnation of advancement in the 21st century. This has occurred mainly as a result of challenges in the translation of benchside basic science discoveries to bedside diagnostics and therapeutics. In fact, since 1995, the US Food and Drug Administration (FDA) has observed a decrease in drug and biologic product submissions for regulatory approval [1]. For example, the number of new molecular entities that have been submitted to the FDA for approval has declined from 45 in 1996 to just 23 in 2010 [2]. This unexpected decrease in the number of innovative medical diagnostics and therapeutics reaching patients has been termed the *pipeline problem*. This problem is a result of the increasingly challenging, inefficient and costly path of medical product development. For example, the time required to bring a new drug to market can be up to 12 years at a cost of hundreds of millions of dollars. The FDA has identified the main cause of the difficulties in the path to medical device marketing approval to be the failure of the applied sciences which are necessary for medical product development in keeping pace with the advances in the basic sciences. In response, multiple government agencies including the FDA, the National Institutes of Health (NIH) and the National Cancer Institute (NCI), along with other public and private partners, have developed an array of initiatives and consortiums with the ostensible goal of tackling this pipeline problem (Table 1) [1, 3–8].

A common theme among all these efforts is the importance of imaging for the future advancement of medicine.

---

J. W. Prescott (✉)  
College of Medicine, The Ohio State University,  
60 King Ave,  
Columbus, OH 43201, USA  
e-mail: jwprescott@gmail.com

**Table 1** Examples of US government initiatives created to address challenges in delivery of innovative medical discoveries

Project name	Purpose statement	Website
Critical Path Initiative (CPI)	“FDA's national strategy to drive innovation in the scientific processes through which medical products are developed, evaluated, and manufactured”	<a href="http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/default.htm">http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/default.htm</a>
Roadmap	“The NIH Roadmap is an integrated vision to deepen our understanding of biology, stimulate interdisciplinary research teams, and reshape clinical research to accelerate medical discovery and improve people's health”	<a href="http://nihroadmap.nih.gov/">http://nihroadmap.nih.gov/</a>
Clinical Trials Transformation Initiative (CTTI)	“To identify practices that through broad adoption will increase the quality and efficiency of clinical trials”	<a href="https://www.trialstransformation.org/">https://www.trialstransformation.org/</a>

Medical imaging provides the ability to detect and localize many changes that are important to determine whether a disease is present or a therapy is effective, by depicting alterations in anatomic, physiologic, biochemical or molecular processes [9]. Quantitative imaging biomarkers are sensitive, specific, accurate and reproducible imaging measures of these changes which can be used in disease diagnosis, treatment planning and medical product development. The use of computer-assisted image processing and analysis is extremely important for the discovery, characterization, validation and application of these quantitative imaging biomarkers. In particular, the application of advanced image processing and analysis procedures (e.g., semi-automated or automated segmentation and registration) provide for more robust and efficient incorporation of quantitative imaging biomarkers into clinical and preclinical decision making.

The objectives of this article are to: (1) introduce concepts of quantitative imaging biomarkers; (2) discuss important considerations in the validation of quantitative imaging biomarkers; (3) describe the use of advanced image processing and analysis for quantitative imaging biomarker discovery and characterization, with an emphasis on opportunities for their utilization in osteoarthritis (OA), Alzheimer's disease (AD) and cancer; (4) discuss current challenges in quantitative imaging biomarker research and (5) present a conceptual framework to use when utilizing image processing and analysis in the development of quantitative imaging biomarkers, which can help to meet some of the research challenges. Overall, this article is meant to promote a common understanding of quantitative imaging biomarkers between image processing and analysis researchers and clinicians or basic scientists and provide suggestions to help advance research in quantitative imaging biomarkers and address the challenges of medicine in the new century.

## Review

### Quantitative Imaging Biomarkers

A *biological marker*, or *biomarker*, is a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or a response to a therapeutic intervention [10]. Biomarkers can be used for diagnosis, prognosis, staging or prediction and monitoring of the clinical response of a disease to a therapeutic intervention. For the case of predicting or monitoring the clinical response to therapy, a biomarker must be associated with a clinical endpoint. A *clinical endpoint* is “a characteristic or variable that reflects how a patient feels, functions, or survives” [10]. Clinical endpoints generally fall under the clinical terms *morbidity* (an incidence of ill health such as stroke or recurrence of cancer) and *mortality* (death). Three difficulties with using a traditional clinical endpoint in clinical studies are that: (1) it may be difficult to standardize or quantify, (2) it may take a long time to manifest and (3) it may be very costly, particularly when a long-term endpoint, such as mortality, is used. In order to expedite the process of clinical analysis, the identification and application of so-called surrogate endpoints is highly desired. A *surrogate endpoint* is “a biomarker that is intended to substitute for a clinical endpoint. A surrogate endpoint is expected to predict clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence” [10]. The utility of surrogate endpoints is a result of their ability to be measured earlier, more conveniently or more frequently than traditional clinical endpoints [11]. Although a surrogate endpoint is one type of biomarker, not all biomarkers are surrogate endpoints; in fact, the large majority of biomarkers are not surrogate endpoints. For example, the World Health Organization (WHO) and Response Evaluation Criteria in Solid Tumors (RECIST) criteria are the only quantitative imaging

biomarkers that are accepted as surrogate endpoints in phase III trials of chemotherapy or radiotherapy of solid tumors. There are no other quantitative imaging biomarkers of solid tumors, such as measurements based on dynamic contrast enhanced (DCE)-MRI or PET imaging, which may be used as substitutes for clinical endpoints [12].

Extending the definition of a biomarker from above, a *quantitative imaging biomarker* can be defined as an *imaged* characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or a response to a therapeutic intervention. Therefore, modalities such as magnetic resonance imaging (MRI), computed tomography (CT), X-ray, ultrasound, positron emission tomography (PET), single photon emission computed tomography (SPECT), optical imaging and light microscopy, among others, can be employed for identification of quantitative imaging biomarkers. Imaging biomarkers can also be semiquantitative, e.g., using a scoring scale based on expert-selected atlases representative of disease severity such as the use of Kellgren–Lawrence (KL) grade for OA [13].

The push towards analysis of medical images for identifying quantitative imaging biomarkers of disease has led to the formation of several projects which mean to develop and disseminate image collections for use in research towards the discovery of quantitative imaging biomarkers. Descriptions of some of these projects are presented in Table 2.

#### Validation of Quantitative Imaging Biomarkers

Quantitative imaging biomarkers are validated by demonstrating an association between the measured biomarker

value and a physiologic, pathophysiologic or therapeutic response. These responses can be manifested by anatomic, physiologic, biochemical or molecular changes. Associations between these changes and disease state can be analyzed using statistical models or classifiers.

There are two criteria that need to be met in order to validate a quantitative imaging biomarker [9]:

- 1) The presence of the quantitative imaging biomarker is closely coupled or linked to the presence of the target disease or condition.
- 2) The detection and quantitative measurement of the quantitative imaging biomarker are accurate, reproducible and feasible over time.

The most stringent validation is required when a quantitative imaging biomarker is being assessed for use as a surrogate endpoint. Weir and Walley have stated the challenge in validating a biomarker as a surrogate endpoint in the following way: “It is insufficient in the validation of a biomarker as a surrogate endpoint to show that it correlates well with the clinical endpoint ... *What is required is that effect of treatment on the biomarker correlates well with treatment effect on the final endpoint*, so that a valid surrogate endpoint allows correct inference to be drawn regarding the effect of an intervention on the true clinical endpoint of interest” [14]. This is effectively an additional criterion for the validation of a quantitative imaging biomarker for use as a surrogate endpoint.

A well-known example of the peril of adoption of an insufficiently validated biomarker for use as a surrogate endpoint occurred in the Cardiac Arrhythmia Suppression Trial [15]. The hypothesis of the trial was that the sup-

**Table 2** Examples of image collections for use in quantitative imaging biomarker discovery and software analysis

Project name	Purpose statement	Website
Lung Image Database Consortium (LIDC)	“...Sharing of lung images, especially low-dose helical CT scans of adults screened for lung cancer, and related technical and clinical data for development and testing of computer-aided cancer screening and diagnosis technology”	<a href="http://imaging.cancer.gov/programsandresources/InformationSystems/LIDC">http://imaging.cancer.gov/programsandresources/InformationSystems/LIDC</a>
Reference Image Database to Evaluate Response (RIDER)	“...A resource for developing and testing computer-aided design software to analyze tumor change at multiple time points as the patient responds to therapy or the disease progresses”	<a href="http://www.nibib.nih.gov/Research/Resources/ImageClinData#RIDER">http://www.nibib.nih.gov/Research/Resources/ImageClinData#RIDER</a>
Alzheimer's Disease Neuroimaging Initiative (ADNI)	“...To define the rate of progress of mild cognitive impairment and Alzheimer's disease, to develop improved methods for clinical trials in this area, and to provide a large database which will improve design of treatment trials”	<a href="http://www.adni-info.org/">http://www.adni-info.org/</a>
Osteoarthritis Initiative (OAI)	“... To develop a public domain research resource to facilitate the scientific evaluation of biomarkers for osteoarthritis as potential surrogate endpoints for disease onset and progression”	<a href="http://oai.epi-ucsf.org/datarelease/">http://oai.epi-ucsf.org/datarelease/</a>

pression of ventricular arrhythmia after myocardial infarction (MI) would reduce the incidence of sudden death. Participating patients were randomly assigned into two groups—those who received an antiarrhythmia drug and those who received a placebo. The suppression of cardiac arrhythmias on an electrocardiogram (ECG) was taken to be a surrogate for the clinical endpoint of reduced mortality after MI. However, it was found that patients who took antiarrhythmia drugs had an increased risk of mortality compared to patients who received placebo. Therefore, the belief that arrhythmias on ECG could serve as a surrogate endpoint for post-MI mortality proved to be false.

#### Improving Quantitative Imaging Biomarkers through the Application of Advanced Image Processing and Analysis

The National Institutes of Standards and Technology (NIST) has identified three primary sources of uncertainty in the use of quantitative imaging biomarkers: (1) the biological variability, (2) the variability associated with clinicians interpreting the images and (3) the physical measurement variability associated with image data collection and analysis across the same or different imaging platforms [16]. Work towards a solution to the biological variability requires the identification of confounding variables within the patient population which may be accounted for before, during or after acquisition and is not particularly amenable to solutions through the application of advanced image processing and analysis. The variability due to clinicians interpreting the images and those associated with image data collection and analysis across the same or different imaging platforms, however, may be decreased by the application of advanced image processing and analysis.

In many cases, image processing and analysis is used to increase accuracy, reproducibility or efficiency of manual analysis. In these cases, an automated processing method may be used to, say, segment a structure of interest. The result of automated segmentation must be compared to those of manual readers as one step in the validation of the process to the application of quantitative imaging biomarker calculation. There are many methods of comparing segmentations, for example, the Dice similarity coefficient (DSC), relative overlap or mean distance [17]. The choice of similarity measure has important implications for the evaluation of calculations derived from a segmentation or analysis as potential quantitative imaging biomarkers. For example, DSC may be used to validate an automated or semiautomated method of segmentation with respect to a manual reader, but if the quantitative imaging biomarker values which are to be calculated from the segmentation are dependent on the surface area of the volume or contour, the use of an overlap-dependent similarity metric such as DSC may not give an accurate measurement and the

resultant quantitative imaging biomarker calculations will be inappropriate.

#### Applications of Advanced Image Processing and Analysis Towards Quantitative Imaging Biomarker Research

Image processing is fundamentally concerned with the extraction of information from an image. The resultant information is then used to derive meaning from the image, a task known as image analysis. The image processing and analysis tasks of enhancement, registration, segmentation and classification play important roles in the consistent and accurate evaluation of quantitative imaging biomarkers.

The possible applications of quantitative imaging biomarkers are broad, but the FDA has defined several diseases in which research into quantitative imaging biomarkers may offer the most benefit in the near future [4, 18, 19]. Three of those diseases are presented in this section: OA, AD and cancer.

In the following subsections, overviews of the use of advanced image processing and analysis for the discovery and characterization of quantitative imaging biomarkers for particular diseases are discussed. Though not intended to be comprehensive, in each subsection, a brief description of the disease, the use of imaging and the anatomical focus of most analyses are provided. In OA, the focus is cartilage morphology; in AD, the focus is hippocampus/medial temporal lobe morphology; and in cancer, the focus is solid tumor morphology and vascularity. Due to the breadth of image processing and analysis methods, in these examples, the methods discussed are focused on the task of segmentation. In addition, due to the substantial amount of different imaging modalities, the modality focused on in the following subsections is MRI. Table 3 provides an analysis of advanced image processing and analysis methods which incorporated quantitative imaging biomarkers for OA and AD.

The literature search for the review of each topic was undertaken through a PubMed search using the keywords “quantitative imaging biomarkers,” “imaging biomarkers” and “computer-assisted diagnosis.” Articles that were included in the review were in two major categories: (1) reviews of imaging biomarkers, especially in the areas of AD, OA and cancer and (2) use of advanced image processing and analysis techniques (e.g., automated segmentation) for the discovery, characterization, validation and application of quantitative imaging biomarkers.

#### *Osteoarthritis (OA)*

OA is a disease that causes degeneration of the cartilage in the joints [20]. It is the most common form of arthritis, affecting approximately 21 million American adults [21]. Clinical onset and progression is generally specified using

**Table 3** Quantitative imaging biomarkers discovered and/or extracted using advanced image processing and analysis methods

Reference	Disease	Delineation	Imaging modality	Imaging biomarker	Method of extraction	Method of validation	Weaknesses
Folkesson <i>et al</i> [32]	OA	Normal vs. OA, represented by Kellgren-Lawrence grade	MRI, 0.18 T, sagittal Turbo 3-D T1-weighted sequence	Medial tibial cartilage volume, medial tibial cartilage surface area	Adjusted k-nearest neighbor classifier, using sequential forward/backward selection to select features of intensity, position, three-jet, eigenvalues and eigenvectors of Hessian and structure tensor, third-order tensor in gradient direction	Inter- and intra-reader variability using DSC; Automatic vs. manual volume measurements using average pairwise difference	No statistical adjustment for known confounders of OA, such as sex, age, and bodymass index; cross-sectional analysis only, no longitudinal data
Williams <i>et al</i> [33]	OA	Change in cartilage morphology in OA patients over a 6 month period	MRI, 1.0 or 1.5 T, with two different sequences: (1) a sagittal fatsuppressed T1-weighted 3-D gradient echo sequence, and (2) a sagittal T2-weighted sequence	Femoral cartilage volume, medial femoral condyle cartilage volume, femoral cartilage thickness, medial femoral condyle cartilage thickness; changes assessed over a six-month period	Anatomically corresponded regional analysis of cartilage (ACRAC), a semi-automated segmentation and atlasgenerating algorithm based on statistical shape models and active appearance models	Intra- and inter-reader coefficient of variation for cartilage volume and thickness measurements	Different MRI field strengths; Study group only included women. No adjustment for confounders of OA, such as age and bodymass index
Morra <i>et al</i> [17]	AD	Normal vs. probable AD, using AD diagnosis and MMSE score	MRI, 1.5 T, 3-D T1-weighted spoiled gradient echo sequence	Hippocampus volume, morphology ("radial atrophy")	Four separate classification techniques, (1) FreeSurfer, (2) SVM with manually selected features, (3) AdaBoost, and (4) SVM with features automatically selected by AdaBoost	Manual vs. automatic segmentation precision, recall, relative overlap, similarity index, Hausdorff distance, and mean distance;	No longitudinal data, therefore cannot assess hippocampal atrophy vs. baseline
Freeborough and Fox [45]	AD	Normal vs. probable AD	MRI, 1.5 T, T1-weighted, axial	Whole-brain texture features using spatial gray-level dependence method	Semi-automatic segmentation of entire brain, calculation of 2D textures using gray-level dependence method, non-linear decision boundary between AD and normal controls tested for statistical significance using Mann-Whitney U test	None; method identifies previously unknown possible imaging biomarkers of AD	No focus on areas of brain most associated with AD changes
Duchesne <i>et al</i> [40]	AD	Probable AD vs. normal controls	MRI, 1.0 T, T1-weighted, fast field echo sequence, sagittal	Jacobian determinants derived from dense deformation fields and scaled image intensity from an automatically selected volume of interest in the medial temporal lobe	Transformation of images into reference space, calculation of deformation field in medial temporal lobe volume of interest with respect to reference space, then class discrimination based on gray-level intensity and Jacobian determinants projected onto reference set eigenspace	Linear discriminant analysis, quadratic discriminant analysis, and support vector machine classifiers for probable AD (clinical diagnosis) vs. normal controls	Volume of interest is a rectangular prism defined in the reference coordinate space, and therefore is nonspecific to the actual shape of anatomical structures of interest (e.g., the hippocampus); Reference data set was acquired with a different magnetic field strength than the analyzed data set.
Gerber <i>et al</i> [41]	AD	Probable AD vs. mild cognitive impairment vs. normal controls, represented by MMSE,	MRI, T1-weighted, from two different datasets: OASIS and ADNI	Coordinates derived from a generative manifold modeling technique	Transformation of images into reference space, then images from subjects with same clinical measure (e.g., probable AD	Linear regression residuals from manifold projections.	Specifics of MR imaging protocol not reported.



**Table 3** (continued)

Reference	Disease	Delineation	Imaging modality	Imaging biomarker	Method of extraction	Method of validation	Weaknesses
		CDR, or clinical diagnosis			diagnosis, MMSE) are used to generate a manifold. Test images are then projected onto the manifold and distances are calculated for subject classification		
Lei <i>et al</i> [42]	AD	Probable AD vs. normal controls, represented by CDR	MRI, 1.5 T, 3-D magnetization prepared rapid gradient echo sequence	Geodesics and initial momenta from large-deformation diffeomorphic metric mapping of the left and right hippocampi	Large-deformation diffeomorphic metric mapping from subject images to template images, from which geodesics and initial momenta are calculated in the template coordinate space, followed by PCA	Automated vs. manual segmentations using an L1 error measure based on the cost of mislabeling voxels	No longitudinal data assessment
Sabuncu <i>et al</i> [43]	AD	Probable AD vs. normal controls, represented by CDR	MRI, from OASIS dataset	Whole-brain template generated from image population analysis	iCluster, a parameterized, nonlinear transformation model which generates template images which represent modes of a set of images; Multinomial logistic regression model performed on iCluster assignments, with regressors for age, education, sex, and MMSE, which demonstrated a statistically significant relationship between MMSE and iCluster assignment	None; method identifies previously unknown possible imaging biomarkers of AD represented as image templates, which correspond to modes of a population	No longitudinal data assessment; no specific focus on structures associated with AD, such as the hippocampus

MMSE minimal state examination, CDR clinical dementia rating

semiquantitative grading schemes of OA severity based on radiographs; these grading schemes are also accepted for regulatory approval of drugs meant to treat OA [13, 22, 23]. The challenge with using radiographs for assessing the entire joint is the poor soft-tissue contrast. Radiographic measurements used for OA grading, such as joint space width, do not provide adequate information about the state of all joint components (e.g., cartilage, bone, muscle, meniscus, etc.). Therefore, newer imaging modalities, such as MRI, are being used more frequently for research into OA. The improvement in MRI technology, including the increased clinical use of high strength magnets, has allowed for the development of image acquisition protocols which can highlight particular joint components [22, 24]. Other factors that have moved research forward in arthritis have been the development of a standard nomenclature for defining joint compartments in the knee, in order to promote consistent research into morphological measurements of cartilage [25], and methods of improving the reproducibility of cartilage thickness measurements [26]. In the case of cartilage, these developments have led to several validated morphological measures—such as volume and thickness—which can be used to assess OA onset and progression (Fig. 1) [27]. In addition, research initiatives, such as the Osteoarthritis Initiative (OAI), have developed a set of standardized protocols which have been used on a large patient population, which allows for the consistent postacquisition analysis of images of patients with OA [28]. Currently, the majority of imaging research into OA is focused on the

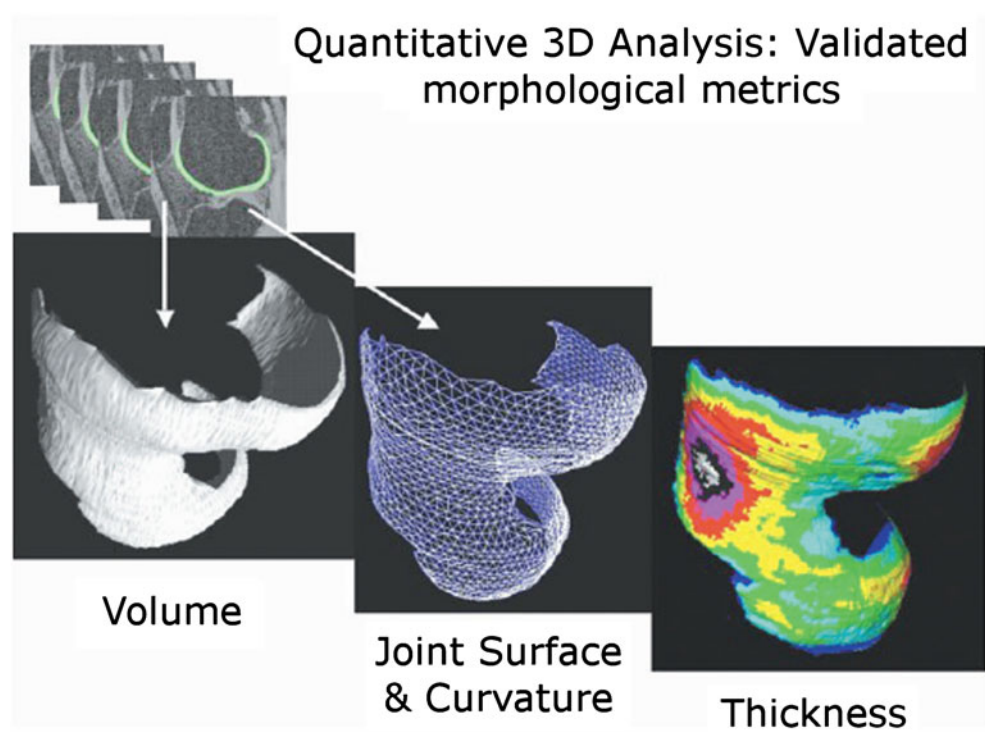
cartilage, but increasingly, there is more interest in a whole-joint approach to the understanding of OA.

There have been several attempts at the automated segmentation of the articular cartilage. Many of these techniques had limitations that prevented analysis of underlying pathology and, therefore, did not extend to an analysis of potential quantitative imaging biomarkers of OA [29–31]. In contrast, some studies have incorporated an assessment of differences between normal and pathologic (i.e., OA) knees, and these studies can be taken as working towards developing quantitative imaging biomarkers for OA onset and progression [32, 33]. Table 3 presents descriptions of these methods.

#### *Alzheimer's Disease (AD)*

AD is a type of cortical, neurodegenerative dementia [34]. It accounts for approximately 70% of all cases of dementia in the elderly, is estimated to occur in up to 30% of adults older than 85 years of age and currently affects nearly 5 million people in the US [34]. Established primary clinical trial endpoints used for AD research are based on symptoms such as functional or cognitive impairment. However, there have been some imaging measures based on structural MRI which have been used as secondary clinical trial endpoints, but none of these imaging biomarkers has been sufficiently validated for use as a surrogate endpoint. Currently, it is thought that the use of neuroimaging in screening at-risk populations (e.g., those over 60 years old) for risk of

**Fig. 1** Validated methods for quantifying cartilage volume, joint morphology and cartilage thickness from three-dimensional datasets have been reported [27]



cognitive decline may be feasible over a time frame of 3–4 years [35]. One goal of the Alzheimer's Disease Neuroimaging Initiative (ADNI) is to characterize morphological changes in the brain of patients with AD [36]. The protocol developed by the ADNI for MR imaging was specifically developed to work with major image analysis methods used in neuroimaging [37].

MR-based morphological measures have been shown to be sensitive in the detection of AD early in the disease course and for the tracking of disease progression. In fact, structural MRI has been shown to be accurate for the detection of disease prior to the onset of symptoms [37]. The medial temporal lobe, specifically the hippocampus, is an area of the brain responsible for episodic memory function in which early AD neurofibrillary pathology is manifested through atrophy (i.e., neuron and tissue lost) [38]. The most common MRI-derived morphological measure in the analysis of AD is hippocampus volume [36]. Recently, however, calculations based on the morphological and intensity characteristics of brain regions in patients with AD have received an increased amount of attention (Fig. 2) [39–45]. Table 3 describes these methods in detail.

### Cancer

The types of cancer amenable to structural or functional image processing and analysis are generally solid tumors; however, there are many different types of solid malignant neoplasms. Although beyond the scope of this article, there are some general similarities in the processing and analysis of images of different types of solid tumors. The FDA has stated that the most valuable role for quantitative imaging biomarkers in cancers is in the use as a clinical diagnostic, rather than as a surrogate endpoint [36]. O'Connor et al. provide a thorough review of the state of biomarkers in the development of cancer therapeutics [12].

In general, morphological measurements, such as length, area and volume, are the most recognized and accepted quantitative imaging biomarkers used in cancer studies. For instance, uni- or bidimensional linear measurements,

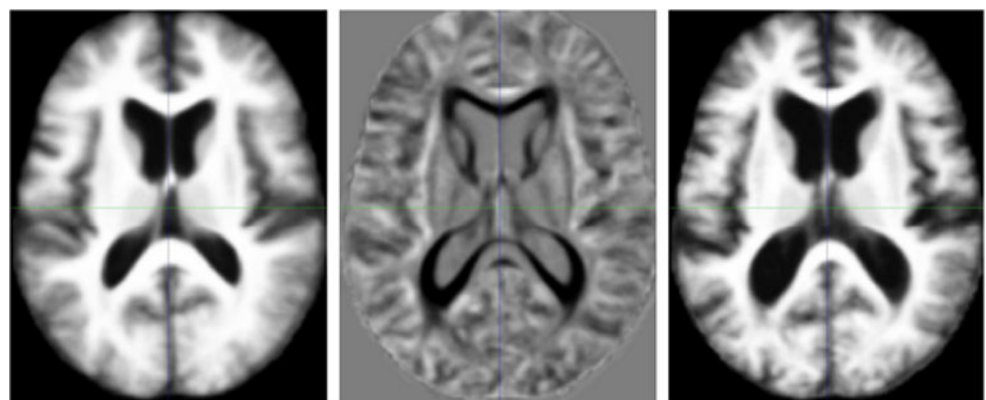
such as RECIST or WHO criteria, are the standard for assessing tumor burden and response to therapeutic intervention in certain studies [46, 47]. However, there have been attempts to employ advanced morphometry, such as contour analysis, in the detection of malignant lung nodules [47, 48]. In fact, the use of computer-assisted diagnosis/detection (CAD) for lung nodules has been one of the most active areas of quantitative imaging biomarker research; however, in the case of CAD, the older term “image feature,” in many cases, may be used as a synonym for quantitative imaging biomarker (although not necessarily always synonymous) [49]. Currently, however, RECIST measurements are the only validated surrogate endpoint in lung cancer [47].

There are a significant number of structural imaging studies analyzing the use of advanced image analysis measurements for the classification of tumors as benign or malignant, for example, the detection of microcalcifications in breast mammograms [50–52] or the quantification of tumor vascularity in a physiologically meaningful way using DCE-MRI through the use of kinetic modeling [53–56] and multi-scale modeling [57]. However, there are new quantitative imaging biomarkers such as the disappearance of intratumoral arterial enhancement, which are being assessed for tracking treatment response [58, 59]. In the future, molecular imaging methods also promise to be an important tool in the analysis of cancer [48].

### Discussion

The path of quantitative imaging biomarker discovery, characterization, validation and, ultimately, application is challenging. However, whereas the criteria required for the validation of a quantitative imaging biomarker (“Validation of Quantitative Imaging Biomarkers”) are stringent, they are not intractable. An important component in the development of any method for extracting and measuring a quantitative imaging biomarker is the early and consistent incorporation of preclinical or clinical information. The developed methods are then required to be consistent and accurate in the calculation

**Fig. 2** Two templates and their difference image from analysis of a 30-patient subset of the OASIS dementia dataset demonstrating differences between patients with dementia and healthy controls [43]





of the quantitative imaging biomarkers that are employed for subsequent preclinical or clinical decision making. The fulfillment of these requirements can be greatly assisted through the use of advanced image processing and analysis techniques, which can reduce the variability associated with different researchers or clinicians interpreting the images and the physical measurement variability associated with image data collection and analysis across the same or different imaging platforms. In this way, advanced image processing and analysis techniques can be leveraged to assist in the resolution of the pipeline problem described by the FDA.

There are several challenges in the incorporation of advanced image processing and analysis into quantitative imaging biomarker research. In Table 3, examples of quantitative imaging biomarkers are presented. The final column of the table lists weaknesses with the validation of the methods in the extraction and calculation of the quantitative imaging biomarkers. In addition to the weaknesses presented, there are general weaknesses within advanced image processing and analysis research when applied to medical images. Following is a list of weaknesses and considerations which can help to address them.

- 1) No healthy vs. disease comparison of algorithm performance: While there are situations where an algorithm may only be required to function on normal or disease images from a single patient (e.g., tracking a tumor's size through time), these situations are in the minority when it comes to possible applications of quantitative imaging biomarkers. Most applications require a comparison of an individual vs. a population, in some way. The most obvious is whether an individual has characteristics of a disease, and those characteristics are based on similar individuals with or without the disease. Fundamentally, medicine works best with dichotomies—healthy or disease? Responding or not responding? Etc. Diagnosis, prognosis and treatment may be based on a continuum, but even these continuums, in general, consist of thresholds and identifiable yes/no questions. Advanced image processing and analysis performed on medical images should take into account where these important dichotomies exist in any current application and perform experiments to validate the algorithm accordingly.
- 2) No adjustment for confounders in statistical analysis: When advanced image processing and analysis researchers seek to define a quantitative imaging biomarker, it is important to include knowledge of the current literature about the disease in question to the statistical validation of the biomarker. For example, it is important to adjust for known confounders of the disease. In the case of OA, known confounders include patient characteristics such as sex, age and BMI, among others. In the case of the analysis by Folkesson et al. [32], the application of their cartilage segmentation algorithm to the classification of normal vs. OA knees is weakened by the failure to adjust for these known confounders in their analysis.
- 3) No longitudinal data in analyses: An important component of any disease state is time; it is an especially important component in chronic diseases such as OA and AD. While a cross-sectional study may be the only type of study feasible given time, patient or monetary constraints, it is nonetheless important to develop a plan or method to expand any analysis to include longitudinal data. The OAI and ADNI provide longitudinal data for OA and AD, respectively. They offer examples of imaging protocols which were developed and implemented with a specific goal of accurate and reproducible acquisition over time.
- 4) Use of inappropriate imaging protocol for structure being processed and analyzed: In some cases, advanced image processing and analysis is part of the discovery phase of research into quantitative imaging biomarkers. Therefore, it may not yet be known what, if any, descriptive measures of disease can be extracted from the given images. However, it is often known which sequence can best delineate (visually, at least) the boundary between a region of interest and “everything else.” By incorporating segmentation and registration together, the performance of both the algorithm for extraction of a structure and the calculation of a quantitative imaging biomarker can be improved. For example, incorporation of “segmentation” imaging protocols (e.g., anatomical sequences) which are quick but accurate with structural delineation may be necessary for consistent and robust application of advanced image processing and analysis techniques. Then, registration of a different set of images which were acquired using a more appropriate protocol for the structure being analyzed can be performed, with the segmentations from the anatomical protocol registered using the calculated transformation. This is, of course, the thought behind the development of registration algorithms, such as maximization of mutual information, and acquisition methods, such as PET/CT. However, there is still room for advancement in the application of these methods to quantitative imaging biomarker research.

With these challenges in mind, it is important to consider a general structure for the incorporation of advanced image processing and analysis methods into the development of quantitative imaging biomarkers for use in clinical and preclinical decision making. The stages of development for quantitative imaging biomarkers are well represented by the five-phase model for cancer screening biomarkers presented in Ref. [60]. In Table 4, the application of that model to quantitative imaging biomarker development is presented. An important thought from this conceptual model is that when considering the application of quantitative imaging

**Table 4** The incorporation of quantitative imaging biomarkers and advanced image processing and analysis into a five-phase model for the development of biomarkers for cancer screening. Adapted from Ref. [60]

Phase	Name	Description	Aims	Incorporation of quantitative imaging biomarkers
I	Preclinical exploratory studies	Exploratory studies to identify characteristics unique to tumor tissue that might lead to ideas for clinical tests for detecting cancer (i.e., comparing tumor with non-tumor tissue).	To identify and prioritize potentially useful biomarkers.	Image acquisition, manual or automated image processing and analysis. In general, measurements of potential quantitative imaging biomarkers would be based on structures of interest delineated by an expert. Features that have strong associations with disease state could be labeled as potential quantitative imaging biomarkers.
II	Clinical assay development for clinical disease	Development of a clinical assay which can distinguish between subjects with cancer and subjects without cancer (e.g., an immune response to a protein uniquely expressed by the tumor and measured with serum antibodies).	To estimate performance of clinical biomarker assay (e.g., true positive rate, false positive rate, etc.) and assess its ability to distinguish subjects with cancer from subjects without cancer.	Development of accurate and efficient methods of calculating possible quantitative imaging biomarkers. This could include automated processing and analysis of the images. The automated methods would need to be validated not only for accuracy and reproducibility against a manual reader, but also for accuracy and reproducibility of calculations of the quantitative imaging biomarkers produced from the automated processes.
III	Retrospective longitudinal repository studies	Compare biomarker measurements from clinical specimens collected from cancer case subjects acquired before their cancer diagnosis with biomarker measurements from control subjects (i.e., those who did not develop cancer).	To evaluate the capacity of the biomarker to detect preclinical disease and to define criteria for a positive screening test in preparation for Phase IV.	Validation of quantitative imaging biomarker and method of calculation for use in detection of preclinical disease. Incorporation of more data into testing of accuracy and reproducibility of advanced image processing and analysis methods used for extraction of quantitative imaging biomarkers.
IV	Prospective screening studies	The screening biomarker is applied prospectively to individuals and definitive diagnostic procedures are applied to those screening positive.	To determine the operating characteristics of the biomarker-based screening test in a relevant population by determining the detection rate and the false referral rate.	Incorporation of more data into testing of accuracy and reproducibility of advanced image processing and analysis methods used for extraction of quantitative imaging biomarkers.
V	Cancer control studies	Addresses whether the screening reduces the burden of cancer in the population.	To estimate the reductions in cancer mortality afforded by the screening test.	Analyze whether quantitative imaging biomarker and method of extraction reduces the burden of cancer in the population.

biomarkers to questions of preclinical or clinical interest, it may be useful to think of the development of a quantitative imaging assay using advanced image processing and analysis.

## Summary

The advancement of medical imaging technology, including an increase in the use of functional and molecular imaging,

provides a wealth of new information to be used in preclinical and clinical decision making. Quantitative imaging biomarkers extracted from the available imaging modalities have become increasingly important in medical product development, disease diagnosis and treatment planning. The forecast is that the role of quantitative imaging biomarkers will only continue to increase. Methods of advanced image processing and analysis can provide powerful tools for the discovery and characterization of

quantitative imaging biomarkers. Future advancements in medical science and industry are likely to benefit significantly from the increased availability of quantitative imaging biomarkers as well as theoretical and software tools for the processing and analysis of the images from which the biomarkers are being extracted.

**Acknowledgments** The author would like to thank Dr. Joseph Yu and Dr. Thomas Best at The Ohio State University for helpful comments and suggestions and Dr. Joanne Lynn and Dr. Catherine Lucey at The Ohio State University College of Medicine for providing funding for the rights to republish Fig. 1.

## References

1. US Food and Drug Administration: Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products, 2004
2. Administration USFaD: FDA New Molecular Entities, 2011
3. US Food and Drug Administration: Innovation or Stagnation: Critical Path Opportunities Report, 2006
4. US Food and Drug Administration: Innovation or stagnation: Critical Path Opportunities List, 2006
5. US Food and Drug Administration: The Sentinel Initiative: National Strategy for Monitoring Medical Product Safety, 2008
6. Administration USFaD: Advancing Regulatory Science for Public Health: A Framework for FDA's Regulatory Science Initiative, 2010
7. Clinical Trials Transformation Initiative. Available at <https://www.trialstransformation.org/>. Accessed December 3 2010.
8. NIH Roadmap. Available at <http://nihroadmap.nih.gov/>. Accessed December 3 2010.
9. Smith JJ, Sorensen AG, Thrall JH: Biomarkers in imaging: realizing radiology's future. *Radiology* 227:633–638, 2003
10. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 69:89–95, 2001
11. Burzykowski T, Molenberghs G, Buyse M: *The Evaluation of Surrogate Endpoints*. Springer Science + Business Media, Inc., New York, 2005
12. O'Connor JPB, Jackson A, Asselin M-C, Buckley DL, Parker GJM, Jayson GC: Quantitative imaging biomarkers in the clinical development of targeted therapeutics: current and future perspectives. *The Lancet Oncology* 9:766–776, 2008
13. Altman RD, Gold GE: *Atlas of individual radiographic features in osteoarthritis*, revised. *Osteoarthritis and Cartilage* 15:A1–A56, 2007
14. Weir CJ, Walley RJ: Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Statistics in Medicine* 25:183–203, 2006
15. Echt DS, et al: Mortality and morbidity in patients receiving encainide, flecainide, or placebo. *The Cardiac Arrhythmia Suppression Trial*. *N Engl J Med* 324:781–788, 1991
16. Clarke LP, Sriram RD, Schilling LB: Imaging as a biomarker: standards for change measurements in therapy workshop summary. *Academic Radiology* 15:501–530, 2008
17. Morra JH, Zhuowen T, Apostolova LG, Green AE, Toga AW, Thompson PM: Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation. *Medical Imaging, IEEE Transactions on* 29:30–43, 2010
18. US Food and Drug Administration: Critical Path Opportunities Initiated During 2006, 2007
19. US Food and Drug Administration: Key FDA Critical Path Activities Under Way in 2007, 2007
20. Buckwalter JA: *Orthopaedic Basic Science: Biology and Biomechanics of the Musculoskeletal System*: American Academy of Orthopaedic Surgeons, 2000
21. Felson DT, et al: Osteoarthritis: new insights. Part 1: the disease and its risk factors. *Ann Intern Med* 133:635–646, 2000
22. Eckstein F, Mosher T, Hunter D: Imaging of knee osteoarthritis: data beyond the beauty. *Curr Opin Rheumatol* 19:435–443, 2007
23. The Atlas of Standard Radiographs of Arthritis. *Rheumatology (Oxford)* 44 Suppl 4:iv46–iv72, 2005
24. Eckstein F, Burstein D, Link TM: Quantitative MRI of cartilage and bone: degenerative changes in osteoarthritis. *NMR Biomed* 19:822–854, 2006
25. Eckstein F, et al: Proposal for a nomenclature for magnetic resonance imaging based measures of articular cartilage in osteoarthritis. *Osteoarthritis Cartilage* 14:974–983, 2006
26. Wirth W, Eckstein F: A technique for regional analysis of femorotibial cartilage thickness based on quantitative magnetic resonance imaging. *Medical Imaging, IEEE Transactions on* 27:737–744, 2008
27. Gray ML, Eckstein F, Peterfy C, Dahlberg L, Kim YJ, Sorensen AG: Toward imaging biomarkers for osteoarthritis. *Clin Orthop Relat Res*:S175–181, 2004
28. Peterfy CG, Schneider E, Nevitt M: The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. *Osteoarthritis Cartilage* 16:1433–1441, 2008
29. Fripp J, Crozier S, Warfield SK, Ourselin S: Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. *Medical Imaging, IEEE Transactions on* 29:55–64
30. Dodin P, Pelletier J, Martel-Pelletier J, Abram F: Automatic human knee cartilage segmentation from 3-D magnetic resonance images. *Biomedical Engineering, IEEE Transactions on* 57:2699–2711
31. Yin Y, Zhang X, Williams R, Wu X, Anderson DD, Sonka M: LOGISMOS—layered optimal graph image segmentation of multiple objects and surfaces: cartilage segmentation in the knee joint. *Medical Imaging, IEEE Transactions on* 29:2023–2037, 2010
32. Folkesson J, Dam EB, Olsen OF, Pettersen PC, Christiansen C: Segmenting articular cartilage automatically using a voxel classification approach. *IEEE Trans Med Imaging* 26:106–115, 2007
33. Williams TG, et al.: Anatomically corresponded regional analysis of cartilage in asymptomatic and osteoarthritic knees by statistical shape modelling of the bone. *Medical Imaging, IEEE Transactions on* 29:1541–1559
34. Andreoli TE, Carpenter CCJ, Griggs RC, Benjamin IJ: *Andreoli and Carpenter's Cecil Essentials of Medicine*. Saunders Elsevier, Philadelphia, 2007
35. Holtzman DM, Morris JC, Goate AM: Alzheimer's disease: the challenge of the second century. *Sci Transl Med* 3:77 sr71, 2011
36. Hampel H, et al: Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. *Nat Rev Drug Discov* 9:560–574, 2010
37. Jack CR, et al: The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging* 27:685–691, 2008
38. Kantarci K, Jack Jr, CR: Quantitative magnetic resonance techniques as surrogate markers of Alzheimer's disease. *NeuroRx* 1:196–205, 2004
39. Martin J, Pentland A, Sclaroff S, Kikinis R: Characterization of neuropathological shape deformations. *Pattern analysis and machine intelligence, IEEE Transactions on* 20:97–112, 1998
40. Duchesne S, Caroli A, Geroldi C, Barillot C, Frisoni GB, Collins DL: MRI-based automated computer classification of probable AD

- versus normal controls. *Medical Imaging, IEEE Transactions on* 27:509–520, 2008
41. Gerber S, Tasdizen T, Thomas Fletcher P, Joshi S, Whitaker R: Manifold modeling for brain population analysis. *Medical Image Analysis* 14:643–653, 2010
  42. Lei W, et al: Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the Alzheimer type. *Medical Imaging, IEEE Transactions on* 26:462–470, 2007
  43. Sabuncu MR, Balci SK, Shenton ME, Golland P: Image-driven population analysis through mixture modeling. *IEEE Trans Med Imaging* 28:1473–1487, 2009
  44. Sabuncu MR, Yeo BTT, Van Leemput K, Fischl B, Golland P: A generative model for image segmentation based on label fusion. *Medical Imaging, IEEE Transactions on* 29:1714–1729, 2010
  45. Freeborough PA, Fox NC: MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease. *Medical Imaging, IEEE Transactions on* 17:475–478, 1998
  46. Eisenhauer EA, et al: New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 45:228–247, 2009
  47. Armato 3rd, SG, et al: The Reference Image Database to Evaluate Response to therapy in lung cancer (RIDER) project: a resource for the development of change-analysis software. *Clin Pharmacol Ther* 84:448–456, 2008
  48. Zhao B, Schwartz LH, Larson SM: Imaging surrogates of tumor response to therapy: anatomic and functional biomarkers. *J Nucl Med* 50:239–249, 2009
  49. Girvin F, Ko JP: Pulmonary nodules: detection, assessment, and CAD. *AJR Am J Roentgenol* 191:1057–1069, 2008
  50. Wang TC, Karayiannis NB: Detection of microcalcifications in digital mammograms using wavelets. *Medical Imaging, IEEE Transactions on* 17:498–509, 1998
  51. Jong Kook K, Hyun Wook P: Statistical textural features for detection of microcalcifications in digitized mammograms. *Medical Imaging, IEEE Transactions on* 18:231–238, 1999
  52. Liyang W, Yongyi Y, Nishikawa RM, Yulei J: A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications. *Medical Imaging, IEEE Transactions on* 24:371–380, 2005
  53. Tofts PS: Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J Magn Reson Imaging* 7:91–101, 1997
  54. Tofts PS, et al: Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols. *J Magn Reson Imaging* 10:223–232, 1999
  55. Armitage P, Behrenbruch C, Brady M, Moore N: Extracting and visualizing physiological parameters using dynamic contrast-enhanced magnetic resonance imaging of the breast. *Medical Image Analysis* 9:315–329, 2005
  56. Schmid VJ, Whitcher B, Padhani AR, Guang-Zhong Y: Quantitative analysis of dynamic contrast-enhanced MR images based on Bayesian P-splines. *Medical Imaging, IEEE Transactions on* 28:789–798, 2009
  57. Mescam M, Kretowski M, Bezy-Wendling J: Multiscale model of liver DCE-MRI towards a better understanding of tumor complexity. *Medical Imaging, IEEE Transactions on* 29:699–707, 2010
  58. Sullivan DC, Gatsonis C: Response to treatment series: part 1 and introduction, measuring tumor response—challenges in the era of molecular medicine. *Am J Roentgenol* 197:15–17
  59. Yaghmai V, Miller FH, Rezai P, Benson AB, III, Salem R: Response to treatment series: part 2, tumor response assessment—using new and conventional criteria. *Am J Roentgenol* 197:18–27
  60. Pepe MS, et al: Phases of biomarker development for early detection of cancer. *Journal of the National Cancer Institute* 93:1054–1061, 2001