RIDER Database Resource and Plans for a Public-Private Partnership

Table of Contents

1. RIDER Database:

1.1 Executive Summary: The Reference Image Database to Evaluate Response (RIDER) to therapy in lung cancer began as a highly leveraged and collaborative *pilot* project, initiated in September 2004, by the NCI's Cancer Imaging Program, NCI's Center for Bioinformatics, the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Cancer Prevention and Research Foundation, and with information technology support from the Radiological Society of North America (RSNA). The specific aims and proposed methods to develop this public resource include:

(a) Development and implementation of an NCI caBIG public resource with the following specific aims:

- Transfer and archiving of de-identified DICOM serial CT and, later, PET CT image data and to include metadata and clinical outcome data where possible, from nationally and internationally distributed sites, acquired during the course of a range of different lung cancer drug and radiation therapy trials.
- Public access to the data by end users (researchers, academicians, device and drug industry) for the purpose of optimizing change analysis software tools and benchmarking their performance, and/or other software development and validation purposes; prior to their use in clinical therapy trials.
- Correlation of the results of change analysis tools that have the best benchmark performance with clinical outcome as a longer term goal, exercised in parallel with NCI or other privately funded drug trials, with knowledge that their clinical performance may be clinical protocol and drug specific.

(b) Development of a broad consensus for the design and implementation of this resource by the RIDER steering committee, comprising academic researchers, program staff at NCI, members of caBIG, NIBIB, FDA (CDRH, CDER) and NIST, and implementation of the this resource using the following approaches:

- Develop a validated database that will include methods of annotation and image markup and demonstrate its use for the assessment of change analysis software tools. The performance of software tools will be tested against this database during various stages of case accrual. One approach will include comparing the variance of a range of change-analysis software tools under development by the RIDER consortium members. This variance metric is expected to serve both as a reference estimate of change over time, and as a means to determine the size and case content of the database, i.e., as required to differentiate the performance of different software tools.
- Seek both national and international input into the design and implementation of the database, that includes publications in peer-reviewed scientific journals and presentations at national and international scientific meetings, including RSNA, ACR, SPIE, and AAPM, i.e., to ensure the public resource is are broadly accepted.
- Develop validation and markup software, implemented in the caBIG Cancer Imaging Archive, consistent with NCI caBIG interoperability and open source requirements of the larger caBIG research community, which include metadata analysis as required for clinical decision making.

Building upon the success of this initial pilot phase, it is proposed that RIDER efforts be expanded from a *pilot* project to a *demonstration* project. As a demonstration project, the RIDER database would be expanded beyond a pilot project to include additional image and related metadata from modalities such as X-ray CT and extended to PET-CT as applied to lung cancer. *A second version of this document (RIDER V2) with address PET CT data, to be completed later this year.* This data will be collected from a wide range of therapy trials supported by NCI and by the pharmaceutical industry. One important goal of this effort would be to engage industry partners in the data collection, database design, and implementation to explore if the database could be

useful in accelerating FDA approval and CMS reimbursement of therapeutic decisions made using software tools.

It is further proposed that this demonstration phase of the RIDER project include both continued support from the existing partners, as well as additional support for the private sector through a Public Private Partnership (PPP). Because the RIDER project is not focused on a specific clinical treatment protocol, or a specific drug, there is little or no intellectual property associated with the generation of this public resource. This demonstration project may thus provide an especially good opportunity to develop a successful PPP of interest to and supported by a broad range of stakeholders in this field, including the imaging, information technologies (IT), software and pharmaceutical industries. The PPP would be coordinated by the Foundation for NIH, a non-profit organization chartered by Congress to raise funds and establish public-private partnerships that complement and enhance NIH priorities and activities.

A longer-term goal, which underlies the use of the term "demonstration project" for this next phase of RIDER, is to use the *organizational structure* of the RIDER PPP as a model for additional projects that would employ, and further evaluate, the most highly rated software tools, using data collected from future NCI clinical trials and, more broadly, trials conducted in other NIH institutes and centers. NIBIB is especially interested an effort to engage not only NCI but other NIH ICs where imaging is being used as a clinical measure for the pathology of disease or its progression and treatment effectiveness.

- By leveraging database resources across NIH in this way, in concert with other efforts such as the Biomarkers Consortium now being established with the Foundation for NIH, the development of more standardized methods for image data collection and validation of change analysis software tools can be accelerated. These goals are consistent with a recent trans-NIH BECON BISTI report (June 2004) submitted to the NIH roadmap committee; the potential for additional roadmap support for these resources could thus be explored. http://www.becon.nih.gov/symposium2004.htm. The goals are also consistent with a planned NIST workshop on standards for biomedical imaging: http://usms.nist.gov/workshops/bioimaging.htm

1.2 Progress Report and time lines for future plans: *The following time lines are approximate and are dependent on the number of data collection sites and level of support for annotation and development of standardized methods for software assessment.*

Deliverables Completed: May 2006:

- Accrual of image data into the RIDER database with advanced lung CA patients; currently over a 150 serial CT exams acquired, over 60 cases on the web.
- Initial annotation of 20 cases using the full implementation of the RECIST criteria, with two independent observers.
- Initial implementation of the RSNA MIRC software for image data collection that meets all de-identification and patient confidentiality requirements and permits a user friendly means to collect and archive image data through firewalls and from multiple sites both nationally and internationally. http://mirc.rsna.org/mirc/query
- Integration of the open source MIRC software as part of the NCI caBIG infrastructure. Initial development of the web accessible public resource and query system, with second release implanted April 2006, version 3 due Sept 2006: http://imaging.nci.nih.gov/i3/
- White paper on database design and requirements for imaging protocols, data annotation, markup, image registration and methods to demonstrate the functionality of this database as a means to measure the relative performance of change analysis software tools.
- Leveraged experience of NCI's LIDC-IDRI Public Private Partnership, initiated Sept 2005 (8 Imaging and PACS companies) to ensure industry requirements are met: for this resource. http://www.fnih.org/partners/research_environment/IDRI.shtml
- Engagement of the FDA CDRH: implemented two fellowship positions at CDRH to address to change analysis problem from a statistical framework, funded by NCI and NIBIB, consistent with NCI-FDA IOTF: http://iotftraining.nci.nih.gov/

- NIST support for imaging standards requested for FY 07 and beyond: http://www.nist.gov/public_affairs/factsheet/bioimaging.htm
- RIDER Initiative: complements the recent DHHS Announcement: "New Federal Health Initiative to Improve Cancer Therapy" (.http://www.fda.gov/oc/mous/domestic/FDA-NCI-CMS.html).

Proposed Deliverables: Fall 2006:

- Completion of Version 2 of this document to cover PET_CT database design for lung cancer
- Complete accrual of over 200 serial CT cases, for a range of lung cancer stages.
- Completion of initial annotation using the RECIST criteria for 100 serial cases.
- Extend the MIRC and caBIG IT infrastructure to include the transfer of annotated, marked up data, and meta-data from multiple sites, to populate the RIDER database.
- Develop a consensus on what constitutes acceptable lung cancer outcomes, for the purposes of correlation with change analysis results.
- Develop an initial consensus with industry on the scope of the RIDER project based on the RIDER WP report. We anticipation plans for expansion of the RIDER project to about 1000 CT annotated cases over two years (2006-2008), depending on the level of interest by industry.
- Develop initial agreements with industry to explore development of standardized methods for benchmarking software tools, in partnership with NIST, to be exercised in 2007-2008. http://usms.nist.gov/workshops/bioimaging.htm
- Initiate a RIDER pilot PET-CT image database resource for lung cancer with data collection from NCI and pharma drug trials. See planned NCI SNM workshop: http://interactive.snm.org/index.cfm?PageID=4901&RPID=4912

Proposed Deliverables for Year 1: Fall 2007:

- Complete the annotation and markup of 500 serial CT cases.
- Finalize the web accessible public resource and query system
- Publish all scientific criteria for CT database design and its functionality.
- Continue collection for 250 cases and initiate the annotation of PET-CT data.

Proposed Deliverables for Year 2: Fall 2008:

- Complete the collection and annotation of all CT cases
- Initial efforts to create standards for software tools assessment with NIST
- Complete the collection and annotation of 500 PET CT data
- Initiate efforts to create standards for software assessment for PET-CT, to be completed in early 2009 or later, in collaboration with NIST.

1.3 Clarification of the Goals of the Federal Trans-Agency Oncology Biomarker Qualification Initiative (OBQI) and the RIDER Project

OBQI Initiative:

Two important categories of biomarkers are biological indicators of disease and markers of therapeutic efficacy. They may involve assessment of genomic and proteomic alterations using laboratory methods or *in vivo* biomedical imaging methods. Recent work has shown that biomedical imaging methods, such as X-ray CT, and more recently FDG PET, provide a means for early indication of drug response.

These research advances have recently resulted in a federal inter-agency announcement of a Memorandum of Understanding (MOU) among the FDA, NIH and CMS. These agencies have agreed to collaborate on improving the development of cancer therapies, in part through the use of biomarkers that correlate with clinical outcomes (*DHHS "New Federal Health Initiative to Improve Cancer Therapy*, *http://www.fda.gov/oc/mous/domestic/FDA-NCI-CMS.html*).

The scope of this MOU covers both the development and clinical assessment of biomarkers in the area of oncology. One specific goal of NCI is to identify those biomarkers that provide a means to measure cancer therapy response. Such biomarkers could be considered "qualified" from an FDA perspective, i.e. they could potentially be adopted as clinical assessment tools in clinical trials intended to be submitted to the FDA for regulatory approval. CMS interests include the development of evidence to make informed reimbursement decisions for biomarkers being used in clinical care. *This trans-agency initiative is referred to as the "Oncology Biomarker Qualification Initiative" (OBQI) and complements the goals of the NCI FDA Interagency Oncology Task Force (IOTF). http://iotftraining.nci.nih.gov/*

RIDER) Project:

There are many sources of uncertainty in the use of imaging as a biomarker for the assessment of drug response. For example, biological variability is a factor that is drug, organ, tumor and patient dependent and thus best addressed through carefully designed clinical trials such as those proposed by the OBQI. However there is also measurement variability associated with two other interrelated factors namely: (a) image data collection across different commercial platforms, and (b) uncertainty in the performance of different image analysis software tools employed to measure therapy response. The latter software tools typically involve measurement of change in image-related computer extracted features over time. Both sources of uncertainty often result in an increase in the required number of subjects in and therefore cost of drug trials. The development of standardized methods to physically characterize these two sources of uncertainty would stimulate the development of improved imaging methods and software tools. The approach proposed for RIDER is to evaluate the proposed software tools for change analysis by first testing their relative performance against a validated and standardized reference database. This approach will include standardized data collections across imaging platforms to populate the database, as this will be also a requirement for multi center drug trials. *The final selection of these well-characterized software tools could then be potentially employed in drug trials submitted to the FDA. The RIDER project therefore has a different technical goal and a different time line to the OBQI initiative but will be very complementary.*

Related URL's:

http://imaging.nci.nih.gov/i3/

http://www.fnih.org/partners/research_environment/IDRI.shtml

2. Introduction: Goals of the White Paper

This white paper regarding the RIDER project has been prepared by multiple authors from multiple institutions, together with agencies of the federal government joining in the common cause in the fight against lung cancer. See the list of authors and institutions in section 8. Lung cancer is the number one cause of cancer death in both men and women as summarized in section 3.1. There is increasing recognition that new therapies are needed for this deadly disease, and indeed new therapies have been, and continue to be developed by the pharmaceutical industry, as well as through NIH funded academic research units. Critical to the clinical evaluation, and ultimate FDA approval and CMS reimbursement, of novel effective therapies is that early and accurate tumor response be obtained, which would substantially reduce numbers of subjects in clinical trials, reduce the length of time needed for clinical trials, and the costs of such studies. Restated, software tools should not be a source of uncertainty in assessment of drug therapy, by imaging or other biosensor data sources. Current efforts of response assessment are inadequate for this task as outlined in section 3.2.

The RIDER project will provide solutions to this problem, through the appropriate use and analysis of imaging techniques. This effort will include image acquisition, image registration, image segmentation, image analysis and image-related statistical techniques, initially in relationship to multi-row detector computerized tomography (MDCT), but extending as appropriate to other imaging techniques such as PET-CT, MRI, US and Optical. The important concepts of truthing in imaging will be initially be explored through the use of simulated and phantom data as well as marked up clinical data and related metrics obtained by a consensus process using imaging experts. Consensus methods are critical to the development of standards for software tools performance as outlined in section 5. Substantial relevant image data sets have already been accumulated as part of an initial proof-of-concept pilot as a web accessible resource, as described in section 4. A process model for the accrual of data into this public resource from multiple academic sites, will be similar to that already developed and in place for the Lung Image Database Consortium, and will be adapted for the RIDER project as described in section 5.4.

At the completion of the RIDER development, a publicly accessible database providing "ground truth" of expert reported RECIST (or other methods) markups on multiple scans and associated meta-data will be available for the broader imaging community. This database will assist the further development of computer algorithms to assess relevant change and the benchmarking their performance. The time table is dependent of the final number of collections sites and rate of data accrual, mark up, and creation of standards for software evaluation, but is expected to take 2-3 years. Also provided will be new information on the many dimensions of image-based truth, fundamentally important to the development and evaluation of change analysis software tools, as required to establish standardized methods. The RIDER project in its entirety provides immediate and appropriate interfaces with relevant NCI resources, as well as with other federal government agencies such as the FDA and NIST. The intent is to provide an interface between the imaging and pharmaceutical industries through the Foundation of National Institute of Health (FNIH) in a RIDER PPP, for their considerable expertise and synergy, described in section 5. The intent is to also highlight the importance of creating standardized methods for benchmarking software tools to reduce the sources of uncertainty for the data integration and clinical decision problem, as recommend by a recent BECON-BISTI symposium entitled "Biomedical Informatics for Clinical Decision Support: A Vision for the 21st Century" June 2004 [http://www.becon.nih.gov/symposium2004.htm].

The battle against lung cancer, as well as the public good, requires the degree of cooperative multi-level effort defined by the RIDER project. This white paper should serve to:

(a) Describe the RIDER project aims and progress to date.
(b) Justify the continued funding by NCI as a pilot project for FY06, and

(c) Demonstrate to the device and pharmaceutical industry sectors that the project is feasible over a 2-3 year timeline.
(d) Through the Foundation of the NIH (FNIH), serve as a basis for the establishment of a RIDER PPP by summer-fall of 2006.

The white paper is expected to have a number of different drafts over the next year. It will be initially be shared with NCI, NIBIB, FDA (CDRH, CDER), NIST and the CRPF, and then presented to the FNIH as an initial framework for the planned RIDER Private Public Partnership (PPP). The FNIH will share the different drafts of this document with the device and pharmaceutical industries to request their scientific input, agree on the final scope, and to then explore the implementation the RIDER PPP, after the scientific goals have been well articulated and documented, as described in section 5.7.

3. Clinical Background.

3.1 Background: Lung Cancer. Lung cancer has a 13% overall 5-year survival rate – a survival rate that has not changed significantly in over 20 years. Lung cancer is the most common cause of cancer deaths in both men and in women in the United States, resulting in more cancer deaths than the next four cancers combined, including breast cancer for women. Lung cancer is strongly associated with cigarette smoking, known at least since the Surgeon General's Report of 1954; however public health measures to abolish tobacco exposure have failed. It is therefore imperative that new novel therapeutic agents be quickly introduced, and rapidly evaluated for effectiveness in this disease. This is beginning to happen (with 151 citations in Medline in which lung cancer and therapy appear in the title in 2005 alone) – however, response assessment has failed to keep pace with potential therapeutic advances.

Lung cancer is classified clinically, for therapeutic purposes, into small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Lung cancer is staged according to the TNM system – the staging gives prognostic information and guides therapy. New methods for staging, such as computerized X-ray tomography, and more recently PET scanning in selected individuals, have been introduced over the last 15 years, with all of them leading to significant upstaging, as the sensitivities of the methods have increased. For instance, with the introduction of computerized X-ray tomography, the disease was upstaged on average 13%.

In the late 1970's, NIH supported an examination of screening for lung cancer in at-risk populations using sputum cytology and chest X-rays; these failed to show benefit. Current studies are underway to re-evaluate screening – using improved sputum tests, and using multi-row detector computed tomography scanning (MDCT), and, in some cases bronchoscopy of at-risk groups, utilizing special bronchoscopes to detect subtle differences in airway mucosal fluorescence as an example of other serial data sources being collected. Thus the use of these different biomarkers will require the development of different interoperable databases to permit evaluation of data integration tools for assessment of therapy response in clinical trials, a task being addressed by NCI's caBIG project. *The RIDER database thus must be fully interoperable with caBIG database archive and other NIH data archives.*

The MDCT studies, most notably the Early Lung Cancer Assessment Program (ELCAP), as well as smaller studies such as those conducted in Japan, and at the Mayo Clinic, have provided important information regarding lung cancer detection. The lung cancer can be seen as a nodule (<3cm in diameter) or as a mass (>3cm in diameter) against the air contrast naturally provided by the lung. More recently, the National Lung Screening Trial (NLST), a multi-center study funded through the NIH, and involving 50,000 at-risk smokers (half randomized to chest x-ray, the other to MDCT) has been initiated. However, it is difficult to discriminate the early changes on MDCT from much more common benign processes that also show as lung nodules. This is particularly so in parts of the US where fungal diseases such as histoplasmosis are endemic. It is emerging that

only 2-3% of detected lung nodules are indeed cancer. The current standard of care, once a small (<8mm diameter) is discovered, is to repeat the MDCT in 3 months and subsequently at 3-monthly or 6-monthly intervals to assess growth of the nodule, on the basis that cancers will grow, and benign diseases will not. This process currently means that the value of early detection (such as by the MDCT) is not translated to early treatment of the lung cancers. This is referred to increasingly as the lung cancer paradox - that is, early detection of small lung cancers is possible, but early treatment is not, as described further below. However, there is substantial work being performed to solve the paradox, with increasing precision of trancutaneous needle sampling, and increasing yields from bronchoscopic biopsies with the assistance of 3D computer-generated graphics (such as from the MDCT), magnetic guidance, and magnetic tracking devices, and the increasing use of ultra-thin bronchoscopes. By 2008, once the NLST is finished, the benefits (or not) of early detection by MDCT will be clarified. Preliminary results may be released earlier if appropriate.

Currently the best hope of lung cancer cure is surgical resection of a small peripheral lung nodule – the cancerous nodule. This is currently possible in about 15% of subjects presenting with early stage disease – the other 85% present with later stage disease, or surgery is not possible because of associated, often smoking-related, co-morbidities, such as cardiac disease, or severe chronic obstructive pulmonary disease (COPD). Those not able to receive curative surgery are often treated with chemotherapy, or chemotherapy with associated usually concurrent external beam radiotherapy. Unfortunately, co-morbid disease often means that patients are not tolerant of other therapies, such as standard chemotherapy or radiotherapy. Information about these aspects, as well as profiles of current therapy options offered to patients are currently being collected as part of the CANCORS study - also funded through the NIH. *However, it is important for the RIDER resource to contain data for early cancers where possible so that tools assessment can be addressed before the new early therapies are implemented.*

Treatment protocols vary across the country substantially. Some patients are eligible for group protocol studies such as through the CALGB or ECOG, as well as other similar geographically related groups. However, only a minority of lung cancer patients are enrolled in clinical trials of any sort. National guidelines exist for the therapy of lung cancer through the NCCN; however, the extent to which these are followed is unclear. In spite of some more recently available chemotoxic chemotherapy agents, the non-surgical therapy of lung cancer usually does not extend life by more than a few weeks, although more significant advances are being made in selected patient subgroups. *Because of the variation in treatment and imaging protocols, the RIDER database proposes to harmonize data collection across different imaging platforms where possible, as described in section 5.1, and to include cases involving different therapies, so that the change analysis methods are applied to real world data as described in section 5.2. Thus data will be collected from on going NCI and privately funded clinical trials.*

In some cases external beam radiation therapy is used as a sole therapy. In other cases, radiofrequency ablation through transcutaneous catheters is also being used for local therapy. In the recent past direct intra-arterial injection of chemotherapy agents (into the lesion's feeding artery) have been used with reasonable local control, but no effect on survival – this approach takes advantage of the information that tumor circulation is derived from the bronchial systemic vessels rather than from the low pressure pulmonary circulation. *The RIDER database resource will thus include data from other forms of therapy, such as radiation therapy. However it recognized that radiation therapy will affect surrounding normal tissue, making measurement of changes in tumor volume, or shape more difficult.*

More recently there is increasing interest and excitement in new biological response modifiers for lung cancer therapy. These are generally less toxic agents targeted to affect the tumor blood supply, or other critical pathways in cancer cell growth, differentiation or metastatic processes. The end point of therapy may not be lung cancer "disappearance", but tumor growth may effectively cease. Additionally, subtle changes in the lung cancer CT density, margins, or

other pixel-based features, may signal a useful response status at an early stage in therapy. Other MDCT-derived features, such as blood flow, may also be important measures in the future; especially taking advantage of the bronchial-artery-derived circulation that predominates in lung cancers. Metabolic changes, as measured for instance by positron emission tomography (PET) scanning, may also give additional information, although PET lacks the spatial resolution of MDCT. *Thus the evaluation of several different change analysis features must be considered for the RIDER database and an extension to other modalities such as PET-CT and MRI to add additional functional or molecular features to the change analysis problem as described in section 5.3.*

Importantly, and representing a serious lack in our knowledge base, we have very limited information regarding the 3D anatomical/pathological structure of lung cancers. This information, together with the MDCT correlates, is fundamental to understanding both lung cancer growth, as well as response to therapy. Current estimates from pathologists working with traditional 2D pathology slides are that the cancer cells themselves only occupy between 15 and 85% of the lung nodule volume, the rest of the nodule mass being inflammatory cells, edema, or areas of fibrosis. (Of interest, the inflammatory component in a lung cancer is an independent prognostic factor). Critical to the image-based evaluation of either growth or response to therapy is gaining a much better understanding of the 3D anatomical/pathological structure of lung cancers, nodules as well as masses, and relating that to the extraordinary resolution in 3D of modern MDCT scanners and the pixel-based measures, some of which are mentioned above. This would appear critical to future computer algorithm developments that could then target the lung cancer component of a lung nodule or mass for response to therapy assessment. *If available, 3D nodule/mass anatomical/pathological high resolution information, with the associated 3D MDCT findings, could either be a sub-component of the RIDER public image database for algorithm development or NCI may encourage other avenues of support through investigated funding for this and other complementary research efforts as described in section 7.*

In many epidemiologically based cancer studies, other patient samples, such as blood and urine, are collected for later analysis for some form of marker that might define the cancer better either for diagnoses, for assessing the effects of therapy, or for prognosis. An example in prostate cancer is the prostate specific antigen (PSA) level. No universal marker has, to date, been discovered for lung cancer since lung cancer appears to be a very heterogeneous disease process. However, there is substantial ongoing work in micro array analysis for genomics, in proteomics and in metabolomics, a recently developing screening metabolic assessment. *The RIDER database should thus be interoperable with planned database involving genomics or proteomics, a caBIG requirement*.

3.2 Background: RECIST Criteria: The method previously used to measure tumor response using CT and other modalities is referred to as Response Evaluation Criteria in Solid Tumors (RECIST). The initial research group that proposed the RECIST criteria included European Organization for Research and Treatment of Cancer (EORTC), the NCI of Canada and NIH-NCI. They reviewed assessment methods for tumors visualized by diagnostic imaging. They examined the existing World Health Organization (WHO) method of bi-dimensional cross-products measurement for evaluating response to treatment in solid tumors, and, in the year 2000 published an alternative analytic approach promoting a simpler uni-dimensional, linear measurement. The RECIST criteria updated the 1979-based WHO methodology to account for technological advances in the tomographic imaging of CT and MRI. It specifically focused on specific measurement recommendations more appropriately matched to image acquisition parameters, illustrating them primarily with CT examples. The paper offered a detailed approach to measurement relating to tomographic slice thickness and scan interval, and specified the size and quantity of baseline lesions. Its prescriptions were supplemented by a web-based question and answer inquiry bulletin board to address the exceptional situations often encountered in real-world trial analyses. Promoting a summative linear measurement of a limited number of target tumors, it offered a simple rapid approach demanding minimal effort and presumably reproducible results.

The premise, however, of both WHO and RECIST criteria presumes that tumors are round and change in a simplistically rounded fashion. Published work comparing RECIST with WHO generally focused on the surrogate of best overall response, with only a few methods addressing other imaging endpoints such as time to progression and disease-free survival. As a therapy response measurement procedure, RECIST, like WHO, still maps summative linear data into an established set of four further discrete categories: CR – complete response; PR-partial response; SD- stable disease; and PD-progressive disease. Those categorical bins, however, are coarse with most trial analyses critically pivoting on PR (partial response) defined by a 30% linear sum reduction, and PD (progressive disease) on a 20% increase in tumor dimension. Publications supporting concordance of RECIST with WHO techniques not withstanding, these linear measurements were determined from subjective observer judgment on often indistinct object boundaries. For example, significant variability in the RECIST measures exists between different observers, and the measurement is a relatively coarse one. Other measures, such as volume, shape or pixel-based derived measures from serial CT may therefore be urgently needed to improve the time-to- response decision making, and to reduce observer variation. If the lung cancer volume is mostly inflammation, then linear size change alone may give a false impression of cancer therapy response (the inflammation itself improved, but the cancerous component did not). For example, when newer chemotherapeutic agents (especially statins) and radiotherapy techniques (stereotactic, proton), the change in tumor size may be inadequate to assess response. Because of these inflammatory reactions, the tumors may slightly increase in size initially. Therefore, in many future instances, new therapy may be effective by stopping tumor growth – in this instance a measure of change in tumor growth may be needed – raising the possibility of a short wash-in observational period to assess tumor growth off-therapy – followed by therapy.

In summary, the historical selection of the RECIST as standardized criteria for therapy assessment was exercised on a rather 'ad-hoc basis', but is the current standard used for drug trials submissions to the FDA. A public validated database resource was not available to provide a quantitative basis for selecting a standardized method for software tools selection. *The planned RIDER demonstration project will serve as this public resource. It will include the use of the RECIST as reference criteria to other proposed CT image features, during the case accrual for the RIDER database, as described in section 5.3.*

4. RIDER Pilot Project: Progress Report (2004-2005)

4.1 Goals of RIDER Pilot Project. The RIDER project was initiated in Sept 2004, as a pilot project. Recommendations for its creation was based on earlier NCI and Trans-NIH workshops related to software tools assessment in 2003-04 and a joint CRPF-NCI sponsored workshops in spring of 2004. The RIDER pilot project was highly leveraged by the earlier development of a public resource for the assessment of CAD methods for lung cancer detection and diagnosis, referred to as the Lung Image Database Consortium (LIDC), initiated in 2001, and described in detail on the NCI web page:
http://imaging.cancer.gov/reportsandpublications/reportsandpresentations/firstdataset.

The LIDC was originally designed as a web accessible 'academic' resource, but was later extended to an 'industry resource', through the offices of the FNIH as a PPP, and referred to as the Image Database Resource Initiative (IDRI). The overall structure of the IDRI steering committee (SC) permitted the role of FDA and NIST scientists, as well as industry to be engaged in the design and development of the IDRI public resource, facilitation a similar SC organizational for the proposed RIDER demonstration project.
http://www.fnih.org/partners/research_environment/IDRI.shtml.

The limited goals of this pilot project can be summarized as follows:

(a) Demonstrate the feasibility for the collection of about 200 serial CT data from different national and international sites that have on-going lung cancer drug trials,
(b) Provide an image archive web access to the image data that is DICOM compatible, within a time frame of 12 months as a public resource, but not to include any marked up data as this process is more time intensive.
(c) Provide a user friendly means for transfer of image data to this resource that is DICOM compatible, ideally open source, preserves patient confidentiality and provides a means to track DICOM data elements be included.
(d) Develop a research plan (RIDER White Paper) to address how the RIDER database should be developed as a demonstration project, the scope to include recommendations for data collection, case selection, methods for image mark-up, and statistical methods for assessment of change analysis tools.

The goals of this Pilot project were met in a very timely way, as described in the following sections below.

4.2 Initial Development of the Web Accessible Image Archive (caBIG)

The cancer Biomedical Informatics Grid, or caBIG™, is a voluntary network or grid connecting individuals and institutions to enable the sharing of data and software tools, creating a World Wide Web of cancer research. The goal is to speed the delivery of innovative approaches for the prevention and treatment of cancer. CaBIG™ is being developed under the leadership of the National Cancer Institute's Center for Bioinformatics. Over 800 people from more than 80 organizations are working collaboratively on over 70 projects in a three-year pilot project. CaBIG™ is delivering tools and applications, all freely available to the community and other interested stakeholders including industry.

The National Cancer Imaging Archive (NCIA) is being developed to support in vivo imaging within caBIG. It provides the cancer research community, industry, and academia with access to image archives that can be used for many purposes. For example, these tools may include the development and validation of analytical software that supports lesion detection and classification, quantitative imaging assessment of drug response and data integration tools where appropriate.

The caBIG Imaging Workspace is part of the caBIG archive funded by NCI and is designed as an outreach to the imaging community, and should prove to be useful for the development of standards for software validation (especially mark-up and annotation schema).

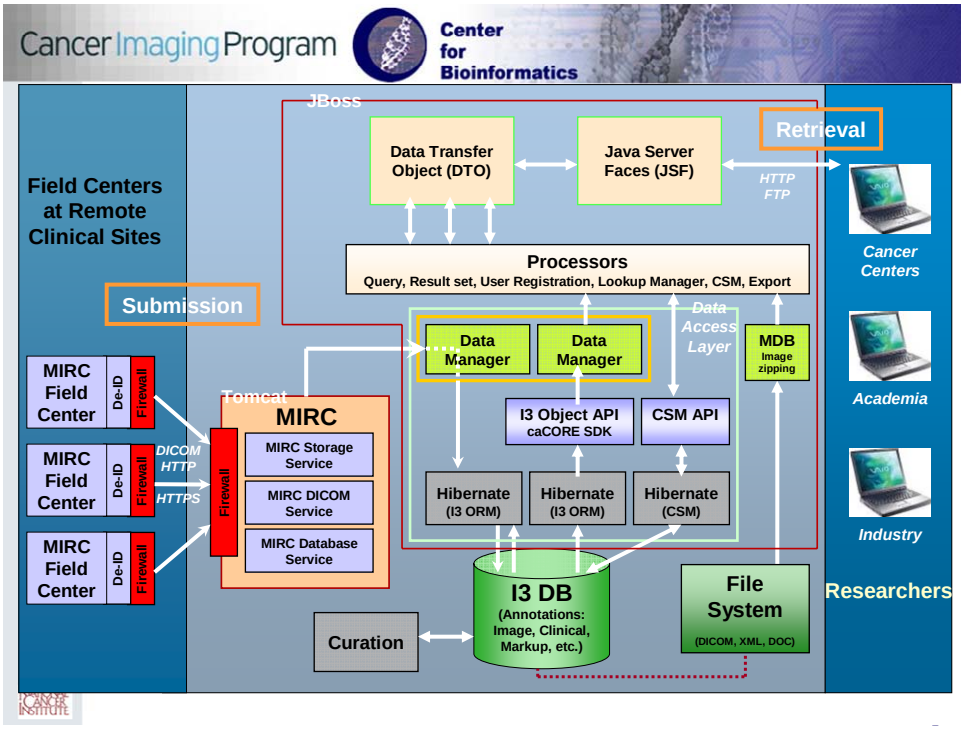


Figure 1: CaBIG and NCIA infrastructure that incorporates RSNA MIRC software for remote data collection that is DICOM compatible.

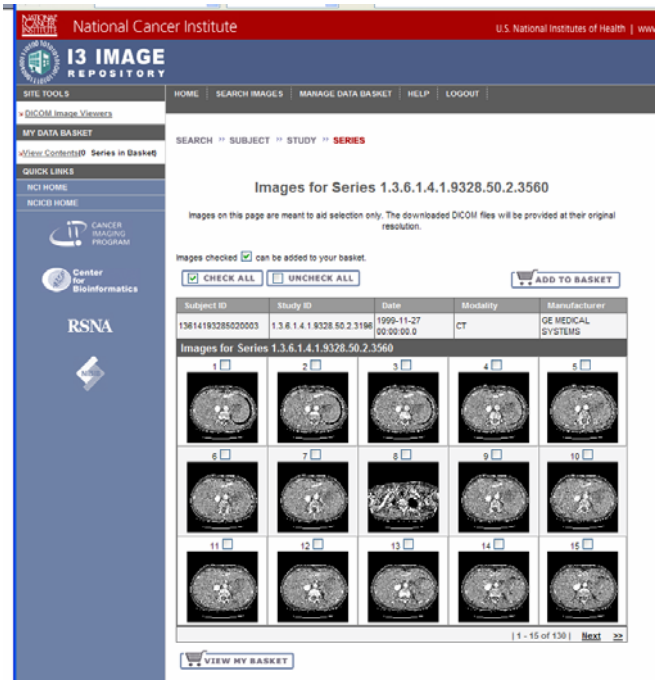


Figure 2: The NCIA website, containing the RIDER data

The NCIA website is designed to provide a scalable and secure archival of images, allowing interactive searching, and interactive downloading and display of images and metadata generated by clinical trials, as shown in figure 2.

4.3 Remote Data Transfer Strategies (RSNA-MIRC)

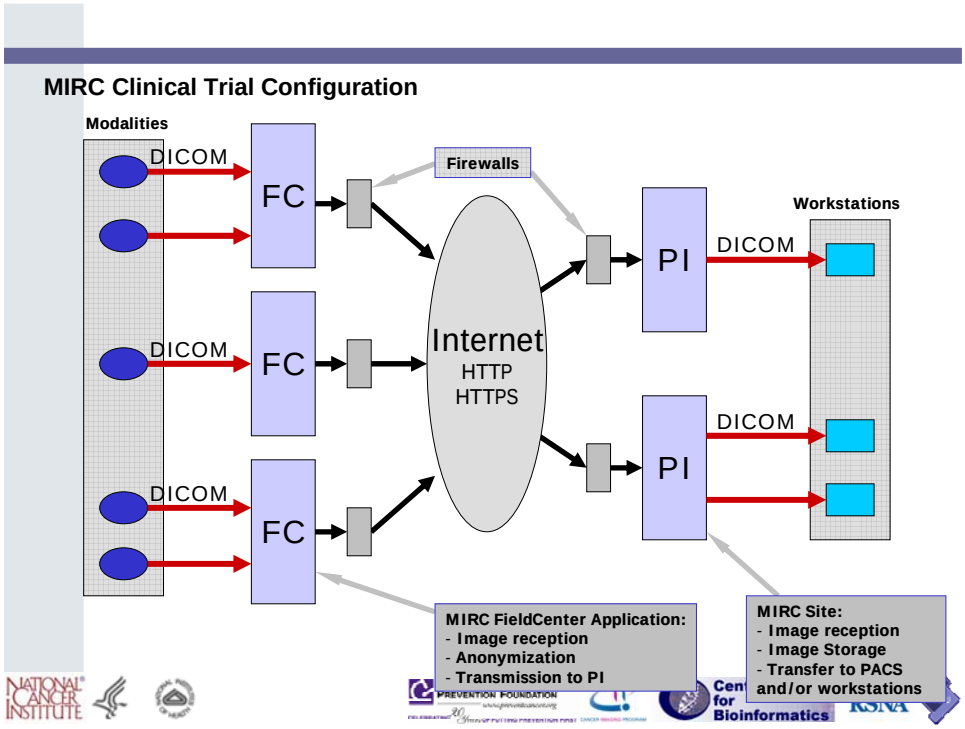


Figure 3. RSNA MIRC infrastructure for clinical trial data collection.

National Cancer Institute's Cancer Imaging Program has selected RSNA's Medical Imaging Resource Center (MIRC) open source clinical trials software suite for use in clinical trials. NCI has signed an MOU with the RSNA during spring of 2005. NCI caBIG staff has since been working with RSNA on setting requirements and refining the MIRC software for remote data collection and data access through different cancer centers, academic and industry sites firewalls. MIRC maintains and tracks patient confidentially. *MIRC is currently only non-proprietary software/standard available free and open source for clinical trials.*

To facilitate the addition of image data to the NCIA, NCI has fully integrated the MIRC server with the NCIA Oracle back – end, and distributes the MIRC field center java program for installation at institutions to provide image acquisition from modalities and PACS, image and metadata anonymization, and secure image and metadata transport, with minimal overhead. *MIRC permits Pharma an effective means of data transfer for prospective clinical drug trials to the NCIA archive and RIDER archive in particular.*

4.4 Initial Progress on Data Collection. Numerous therapeutic clinical trials are conducted by government and private sponsorship that include imaging at several time points before and after therapy, particularly CT, PET-CT and MRI, where accurate tumor measurement may provide useful indication of disease progress. Many trials sponsored by NCI's Cancer Therapy Evaluation Program (CTEP) use network-connected imaging instruments capable of transmitting to

a central archive digital studies in DICOM standard format. Collections of time-sequence patient-studies, de-identified to be HIPAA-compliant, can be made available for computer image processing research. *The RIDER pilot project data was therefore initiated on Sept 31st 2004 where the data was obtained from on-going clinical trials, considerably reducing the cost of the development of this public resource.*

The data collection for this pilot project was initially targeted at 200 serial cases using X-ray MDCT with stage 2 or 3 lung cancer. Each case had a diagnostic CT scan (for staging) followed by multiple time-point sequential exams throughout the course of therapy for a total of 3 to 8 studies per patient. Cases were selected from patients undergoing systemic chemotherapy treatments for late stage disease so that few will have undergone the localized anatomically-distorting effects of radiotherapy or surgery. Cases mounted in the public archive, as shown in section 4.2 (Figure 2) were chosen to exclude major artifacts such as hardware (pace-makers, etc) or breathing-motion artifacts. Initial case selection, however, did not exclude mediastinal or pleural involvement since there will always be a dominant lung mass predominantly surrounded by aerated pulmonary tissue. Target quantification efforts may choose to focus on one of potentially several visible tumors, which may be chosen as representatives of the disease, as is done in the established RECIST method. All cases were collected from clinically-active cancer centers of national or international reputation as listed in section 8. The images were obtained in DICOM format from state-of-the-art multi-detector row CT (MDCT) machines. For retrospective cases, the scan protocol was set at 5 to 7.5 mm slice thickness at adjacent cuts, which is common practice for imaging patients with established lung cancers; these are the thickness parameters for the initial datasets. However, thinner collimation datasets, with thickness ≤ 2.5-3.0 mm desired and ≤2mm preferred, are being included where possible in this pilot database. The images were transferred to the NCI image archive in a de-identified fashion via MIRC transfer protocol (RSNA) described in section 4.3. Clinical outcome and other clinical data have not accompanied the pilot collection but such data may be included in the later RIDER demonstration project, when prospectively specified, as described in section 5.3.

At present about 120 separate lung cancer MDCT image sets (an initial MDCT and then sequential studies during therapy) have been collected and transmitted to the NCI, the target of 200 cases is expected to be met by summer of 2006. An additional 40 cases have been collected from the international sites under the CRPF funded efforts.

At present about 120 separate lung cancer MDCT image sets (an initial MDCT and then sequential studies during therapy) have been collected and transmitted to the NCI, the target of 200 cases is expected to be met by summer of 2006. An additional 40 cases have been collected from the international sites under the CRPF funded efforts.

4.5. RIDER Pilot: Progress on Boundary Identification of Nodules: The development of the NCI-sponsored lung imaging database consortium (LIDC) for testing the relative performance of different algorithms for lung nodule detection and classification, http://imaging.cancer.gov/programsandresources/InformationSystems/LIDC], and the RIDER project as related to measurement of therapy response have similar goals. While the LIDC project requires accurate segmentation of nodules, the RIDER project requires the accurate measurement of nodule change over time. Since the majority of small lung nodules are not resected, a reference standard from histopathology is generally unavailable, and histopathology of nodules may not provide sufficient spatial information. Thus understanding the impact on variability in defining the spatial extent of lung nodules by radiologists using different software tools is necessary.

As part of the LIDC effort, a study investigating the variability of boundaries across radiologists and drawing methods was performed. In this study, the performance of 6 radiologists each applying three different outlining methods to the task of defining the spatial extent of 23 different lung nodules was evaluated. The performance of 6 radiologists each applying three different outlining methods to the task of defining the spatial extent of 23 different lung nodules was evaluated over the last 6 months. The drawing methods consisted of one entirely manual

method, and two semiautomatic drawing methods with manual editing tools. The variability of radiologists' spatial definitions for a nodule was measured using both volumes and probability maps (p-map), where the value of a voxel in the p-map represents a spatially smoothed estimate of the fractional number of radiologist-method combinations that include that specific voxel inside their boundary definition of the nodule. Both volumetric and p-map experimental data were modeled using general linear models that included nested random effects.

Differences in volume and p-map model parameters were found to be significant for all methods, all radiologists and all second order interactions except one across the combination of all 23 nodules. The radiologist and methods variables accounted for 15% and 3.5% of the total p-map variance and 40.4% and 31.1% of the total volume variance, respectively. Thus in the critical process of segmenting lung nodules radiologists demonstrated 4 times the variance of the methods used, and in the related assessment of nodule volumes radiologists demonstrated 1.3 times the variance of the methods. A report on this work has completed and has been submitted to the Journal of Radiology for publication on May 5th 2006.

4.6 Rider Support for Pilot Project for FY 06. The office of the Director of NCI charged the different programs at NCI to establish a unique research plan to improve programmatic interactions within NCI and develop research and related resources to address the cancer problem by 2015. The I2 plan, referred to as the Imaging Integration Plan (I2 Plan) includes a business organization model with deliverables of products. One of the goals is to leverage NCI support with other agencies and the private sector. The I2 plan was funded in late summer of 2005 and provided initial support for the IDRI PPP project. Support for the following FY 06 has been approved by NCI and includes:

(a) Support for database resources such as the RIDER project, with extension to include other database resources such as PET CT imaging to access response to lung cancer drug therapy and other modalities and organ systems in the out years. One goal is to provide continued NCI matching support for the planned RIDER PPP in FY's 06-and beyond.
(b) Support for expansion of the caBIG archive to archive image data and related metadata, query based web search of this data, and development of open source annotation and mark up software and IT tools for both the planned imaging workspace and targeted databases for software tools assessment.

4.7 NCI-NIBIB Supported FDA Fellowships and Visiting Scientists. NCI and NIBIB have recently agreed to support two research fellows that will work under the research guidance of the staff at the FDA CDRH. The research fellowship positions will be initiated by June 2006. The research scope will focus on the statistical methods necessary to measure change analysis over time using biomedical imaging. The initial work will focus on the RIDER project as a collaborative effort by NCI and NIBIB, and will later extend in FY 2007 to other closely related projects such as the NIA-FNIH ADNI PPP project, involving the use of serial MRI and PET imaging as a biomarker for drug response, a project partly funded by NIBIB.
http://www.fnih.org/programs/research_environment/ADNI.shtml.

In addition, the FDA CDRH and NCI program staff have identified several candidates to apply for the Inter Agency Oncology Task Force (IOTF) Fellowship Program for the FY 2006, where the candidate(s) will also explore the change analysis problem in the context of drug response, consistent with the aims of the IOTF and OBQI initiatives, under the direction of FDA and NCI research staff.
http://www.fda.gov/bbs/topics/news/2005/hhs_021605.html.

Finally NCI is now supporting a visiting PET CT physicists-computational scientist for 10 % time as of June 1st 2006 to provide advice on how to standardize data collection and provide advice on the population of the RIDER database. See NCI-SNM workshop below.

http://interactive.snm.org/index.cfm?PageID=4901&RPID=4912

4.8 NIST Collaboration on Imaging Standards. Over this last year NIST has expressed an interest in the development of resources that would serve as a basis for exercising standardized methods for software assessment for biomedical imaging. As a result of this collaboration NIST scientists are represented on the LIDC IDRI and RIDER Steering Committee since summer 2005. NIST is organizing a workshop on the development of physical performance standards for imaging platforms, and for change analysis software tools entitled "Imaging as Biomarker: Standards for Change Measurements in Therapy" (Sept 14-15th 2006). The RIDER project will be highlighted as a model inter-agency initiative for the development of imaging standards. NIST has proposed significant support for this effort in their FY 06 plans to be submitted to congress.
http://www.nist.gov/public_affairs/factsheet/bioimaging.htm

5: Proposed RIDER Demonstration Project

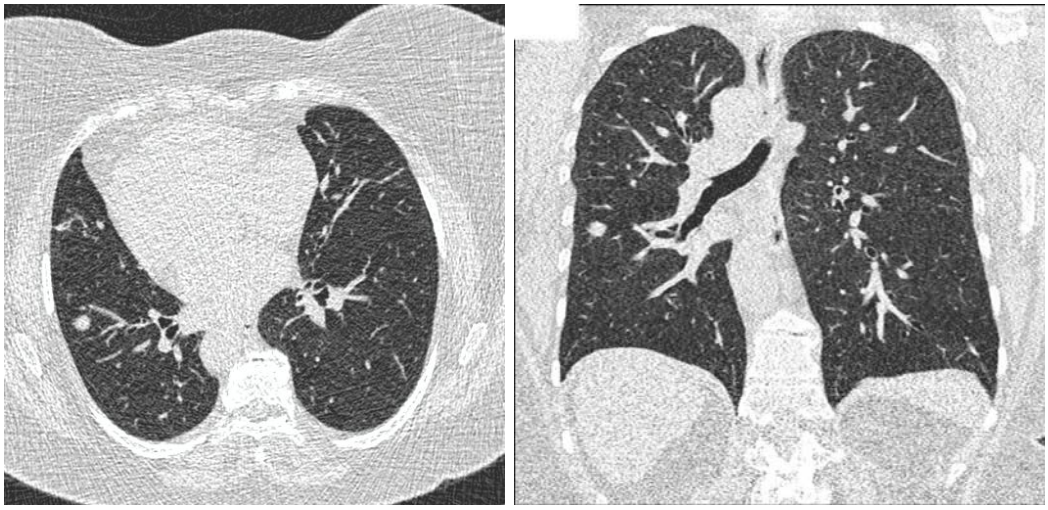
5.1 Imaging Data Collection: Acquisition Parameters. To carry out accurate assessment of change, the database must contain detailed descriptions of acquisition parameters that could potentially influence the accurate measurement of that change. As described in section 3.1, these metrics of change may be tumor morphology (diameter, volume, attenuation, etc.) or other important metrics of change such as functional parameters (active tumor volume, tumor perfusion, etc.). The acquisition parameters to be recorded will therefore vary by imaging modality, but some examples will be provided below. For CT imaging, analysis of the image data should be performed knowing several key imaging parameters: (a) CT scanner make, model, technical capabilities, (b) whether the imaging performed was done as a single static scan (such as a single breath hold scan) or as a sequence of scans that provide a dynamic scan (such as a series of scans performed both before and after the injection of a contrast agent); (c) technical parameters such as the slice thickness, x-ray beam energy in kVp, tube current, exposure time (or tube rotation time), pitch (defined according to the IEC definition of pitch), reconstruction filter, reconstruction interval (e.g. spacing between slices to provide contiguous or overlapping images) and a few others, and (d) information about whether any contrast agent was used in the imaging and if so, then what type, concentration, rate of injection, timing of scan with respect to injection and total amount of contrast (especially with respect to patient weight).

For other imaging modalities, the set of technical parameters will also have to be defined. For example, in PET-CT, the imaging parameters above for CT will have to defined as well as the technical parameters for the PET scan, which will include information about the radiopharmaceutical used (e.g. ^{18}F-FDG or other positron emitter), concentration and dose of the agent, timing of the scan, extent of anatomy imaged, and many other related imaging physics parameters such as the reconstructed slice width, attenuation correction methods (including the use of PET-CT data), gating methods to reduce motion artifacts etc. Also, when image fusion is performed, the algorithm for fusing the PET and CT data should be described. As another example, when MR imaging is performed, the description of whether the imaging is done in either a static breath-hold or continuous breathing or even a dynamic scan condition should be described accurately in addition to technical parameters analogous to the CT scan parameters described above (e.g. T1 weighting, T2 weighting or other sequences, etc.) and information about inhaled (e.g. hyperpolarized Xn gas) or injected contrast agents

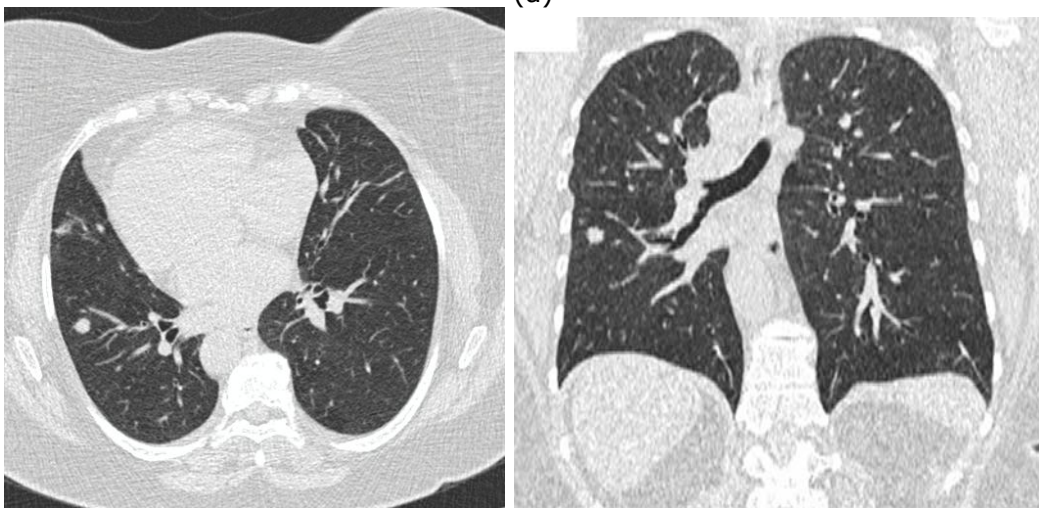
As the RIDER database will initially focus on images from CT, the following technical specifications for inclusion of image data into the database were proposed for initial review by the RIDER project team. These inclusion criteria were primarily focused on slice thickness and not on other technical parameters. These criteria are: (a) slice thickness of $\leq$ 7.5 mm (with thickness of $\leq$ 2.5 mm preferred), (b) single breath-hold scans, (c) patients must have had at least two scans over time and preferably three to assess progression with one of the scans preferred to be at baseline, (d) patient must be on a chemotherapeutic or radiation therapy treatment protocol or a

joint protocol, (e) patient must be under treatment for primary lung cancer (preferred) or have a primary malignancy that is metastatic to the lung, (f) the patient must have some measurable disease on the first scan, but the cancer can be at any stage (of the TNM staging system), (g) and it is expected to include some scans that are performed after contrast injection. All image data provided to the RIDER database were anonymized and all Protected Health Information (PHI) was removed in accordance with HIPAA guidelines. Identifiers will be provided to link images from the same patient together, but these identifiers will not refer to the original patient identifiers.

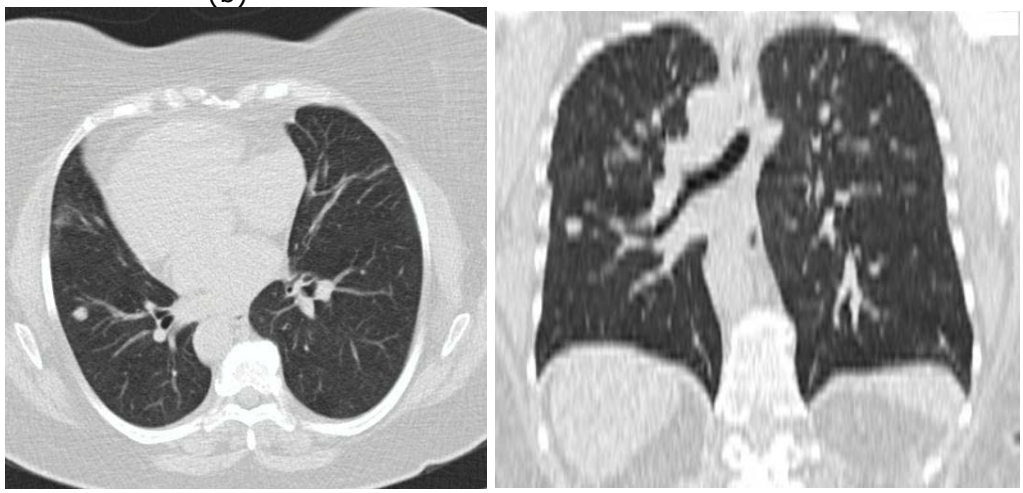
For the RIDER project full demonstration project, we would strongly recommend acquisition protocols in which thin slices (<= 1.25mm thick slices) would be made available for the database. This would allow more accurate estimates of tumor volumes, diameters and other morphological descriptions that may change over time. Collection of data with thin slices should be very possible with the proliferation of 16 (and higher) slice scanners and the ability to perform multiple prospective reconstructions (Note: This could even be done with archiving of the very large raw projection data at the site of the originating scanner. The technical details of this would have to be worked out with each site as the data collection goes forward.) Some example images reconstructed with different slice thickness and different reconstruction filters are included here (see Figures 4 and 5) to illustrate the importance of these technical parameters on image quality, lesion appearance and their potential to affect reliable measurement of tumor volumes.



(a)

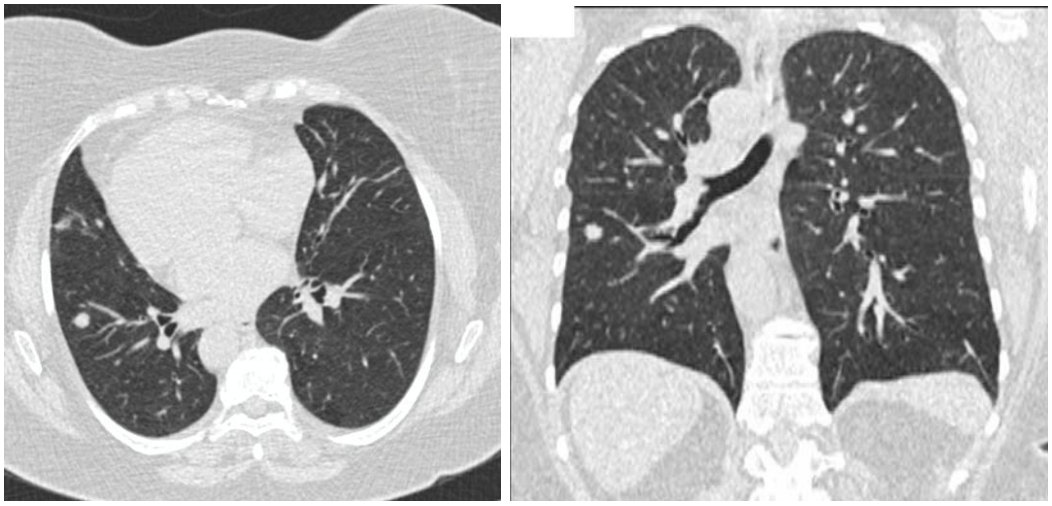


(b)

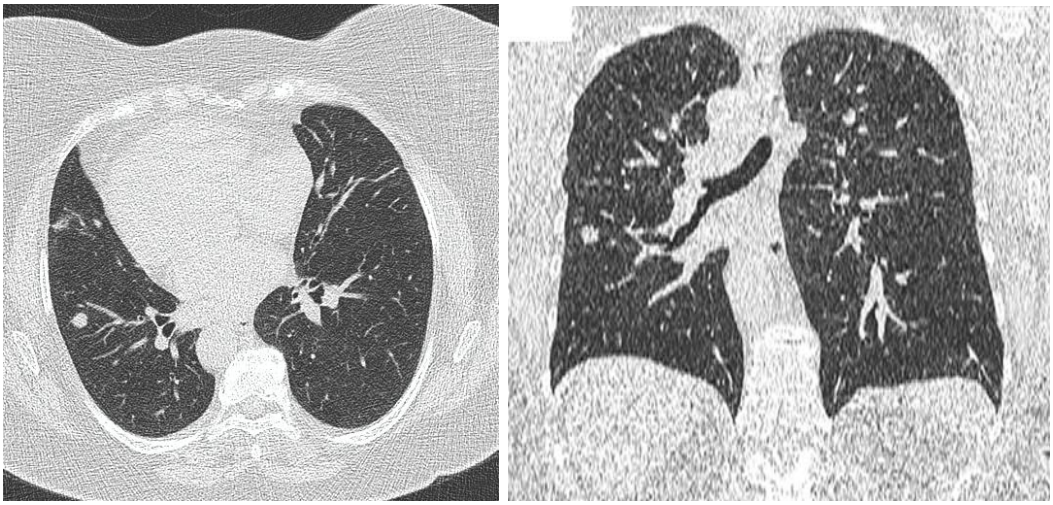


(c)

Figure 4: Figures illustrating differences in image quality and lesion appearance due to differences in slice thickness. This figure shows axial and coronal images of the same patient (in fact, the same original scan acquisition) with reconstructed slice thickness of (a) 0.6mm, (b) 2 mm and (c) 5 mm



(a)



(b)

Figure 5: Figures illustrating differences in image quality and lesion appearance due to differences in reconstruction algorithm (recon filter). This figure shows axial and coronal images of the same patient (again, the same original scan acquisition) having 2mm thickness and (a) smooth reconstruction algorithm, (b) sharp algorithm

5.2 Populating the Database: Case Selection Criteria. The database for the RIDER demonstration project will consist of radiologic images of malignancies. Primary data will include CT images from standard CT equipment and/or CT data from PET/CT. For the initial pilot database, images were collected from the existing inventory of scans that were acquired as part of the sites' clinical operation; these images were typically acquired at full inspiration with thicker collimation. Thus, while initial datasets contained images up to 7.5 mm slice thickness, thinner collimation datasets would be desired in the near future as described above. A representative variety of lesions will include various sizes, locations, and extent to reflect tumors typically encountered in cancer imaging. Ideally, there would be representation of lesions totally isolated within the lung and lesions adjacent to or invading nearby structures such as the chest wall, mediastinum, or heart. Well marginated as well as poorly marginated (spiculated) lesions should also be included. Cases from histologically proven bronchogenic carcinoma would be included, but pulmonary metastatic lesions are also acceptable. Intravenous contrast examinations are recommended, but not required. At least 1 follow-up study is required, but serial imaging is preferred. *Finally case selection clearly will be dependent on the additional collections sites from academia and pharma. Two Pharma sites have agreed to provide PET CT data from prospective drug trials, details as yet unknown.*

Patient information should include standard demographics (sex/age), smoking history, histological tumor type, clinical stage, current and interval therapy, including chemotherapy and/or radiotherapy. Clinical outcome data will be added when available. Rationale for the type of therapy is that with conventional chemotherapy and radiotherapy, the size (volume) assessment may be sufficient to assess response. However, the newer chemotherapeutic agents (especially statins) and radiotherapy techniques (stereotactic, proton) the usage of tumor size may be inadequate to assess response. While it would be good news for the tumors to decrease in size with these newer treatments, in fact, because of inflammatory reactions, the tumors may slightly increase in size initially. Current experience indicates that most tumors do remain stable in size with these newer therapies. Therefore, in the future, it will be important to look at other parameters, such as attenuation changes, margin contour alterations, or pixel based changes over time, especially as better imaging techniques and registration methods are developed. This also means that future databases should include the therapeutic agent.

5.3 RIDER Database Design: Initial research Plans. The results as outlined in section 4.5, coupled with the desire to measure nodule changes in response to therapy in the RIDER project suggest that supervised application of semiautomatic methods of nodule segmentation may provide lower variance estimates of volume change which in turn would allow the reliable detection of smaller changes in nodule volume. In the LIDC project defining truth required the use of methods with low bias potentially at the expense of higher variance, but the nodule change detection task of the RIDER project requires the use of lower variance volume estimation methods potentially at the expense of increased bias, traits which are readily available from semiautomatic methods.

Similar to the drawing experiment described in section 4.5 the RIDER project should compare the sensitivity and accuracy of different methods/tools in detecting nodule volume change. All of the following experiments should be performed by multiple supervisors using multiple methods; additionally multiple repetitions of the same supervisor/method combinations should occur in order to estimate intra- and inter-supervisor effects. It is easily possible that a single supervisor using a single method may have less variance than that measured across different methods; while the latter can be estimated without repetitions of the same method, the former cannot.

Volume change detection methods evaluated across multiple exam series should include:

- The current standard of practice, i.e. the manual RECIST method described in section 3.2 where measurements are in progress.
- Automatic/semiautomatic tumor segmentation methods and methods based on supervised automatic tumor registration techniques including:
    - Local rotate-translate registration followed by gray-scale weighted subtraction, and
    - High degree of freedom warping followed by integration of the Jacobian.

Data acquired to support the evaluation of the above methods for measurement of

- Bias error, i.e. error in the measurement of average accuracy, should include
    - scans of phantoms containing objects of known volume change,
    - scans modified using known mathematical geometric deformations,
    - Variance, i.e. measurement noise which results in the reduction of sensitivity in the detection of a small volume change, should include the following data sets where we are certain that there are no volumetric changes in the nodule,
    - Hourly or shorter interval exams of patients with nodules of varying size and complexity that volunteer for multiple scans,
    - Scans from patients with follow-up exams over several years that demonstrate nodules with no visible change over the history of the patient.

Assuming that the most important goal desired by pharmaceutical companies is to detect early, small nodule volume changes which might predict therapeutic outcomes, the highest priority experiment listed above to perform is the measurement of variance of the different methods when testing on nodules with no [no or little observable] change. This measurement of variance essentially establishes the Student's-t distribution with which to threshold volume change measurements as to whether they are statistically significant, i.e. a true change signal, or insignificant, i.e. most likely a result of the noise of the measurement system. As the estimate of variance is in itself a noisy measurement, a preliminary experiment would be to run a small pilot study of approximately 20 serial cases to roughly approximate the variance of the different methods to be tested. Using these estimates of variance, the participants could then power the follow-on experiment to establish significantly different variances between the proposed methods at a chosen level of confidence. Determining other measures such as bias are less noisy can could be determined in smaller follow-on experiments.

The resources to achieve the above aims can readily be accommodated. The RIDER PI's have agreed to provide their drawing tools for this experiment, and other RIDER research sites and some industry research sites are interested in providing segmentation tools, some of which require minimal observer input, such as defining an area enclosing the nodule and identification of the approximate nodule centroid. The NCI imaging workspace as described in section 4.2 may provide a means to support this experiment.

Timeline: The intent of this continued effort for the RIDER pilot project is to explore how to provide "spatial change truth". With advances in automatic segmentation and registration tools, these tools may be implemented to provide serial estimates of "spatial change truth" for the RIDER demonstration project. This approach would minimize the need to define the boundaries of these nodules manually by expert radiologists. The intent is to be able to show some provisional results before the RIDER demonstration project is initiated as a PPP summer-fall of FY 06.

5.4 Project Multi Site Organization:

5.4.1 Process Model. The collection of data for the RIDER project is being modeled after that developed for the Lung Image Database Consortium (LIDC). In that data collection effort, a process model was developed to describe the steps involved in the collection of image data, the annotation of that image data by expert radiologists and the addition of other demographic and clinical data. While the RIDER project has a different specific aim than the LIDC, that of providing a reference database to evaluate image based response to lung cancer therapies, much of the data collection process from LIDC can be leveraged to create the RIDER database.

For the LIDC, one of the key elements of that database was to provide detailed information about the location and extent of lung nodules based on expert radiologists' interpretation of the image data. However, members of the LIDC had observed in their own research and in what was being presented in the scientific community that there was considerable inter-reader variability in the detection of lung nodules from screening CT exams. Rather than try to provide a consensus-only approach (which was going to prove difficult to perform across institutions), the LIDC decided early on to develop a process that would both provide the best estimate of the radiologists' interpretation of the image data as well as capture any variability among the readers. The result was a data collection process that utilized multiple readers performing two readings of the same case in what the LIDC came to call a "blinded and then unblinded" reading process. In this process, each of multiple readers (four independent readers to start with) read each case independently; that is, they read without knowledge of how any other reader read the case or what each other marked. The results of each independent read was then combined into a set of annotations that could be displayed with the image data and sent back to each reader; in this way, each reader could see not only what they themselves had marked on the image, but they could see – unambiguously – what the other readers had marked in the first pass independent read. In this second read, each reader was now "unblinded" so that they could see what the other readers had marked. Each reader could now not only see what other readers had marked, but they could choose to agree or disagree with the other readers; note that there is still no forced consensus in this process. The final result for each reader was their markings after this second unblinded reading. The results from each reader were then combined to create the unblinded reading result, which represents the best estimate of nodule locations and extent as well as a measure of agreement between readers.

Clearly, the LIDC was not satisfied with traditional textual descriptions of nodules and sought a mechanism to record as closely, and as unambiguously as possible, the radiologists' observations of the image data. Therefore, the LIDC first set guidelines for three general categories of what objects were to be marked in the image data as well as what data was to be recorded for each type of object: (1) For nodules =>3mm, a detailed outline of the nodule was to be provided on all slices on which the nodule appeared; in addition, there were several subjective ratings of the nodule's characteristics that were to be provided (such as the nodule's subtlety, shape, composition, etc.); (2) For nodules < 3mm that are not clearly benign, the LIDC wanted to at least include a marking that indicated the approximate centroid of this object; it was decided that contours of 3mm nodules would not be accurate and that subjective assessments of such small objects would not be helpful, so only the centroid marking is made; (3) Because there are so many nodule-like objects that appear in CT images of the lungs, the LIDC also decided to provide markings of objects that appeared in the image and could potentially confuse users of the database (and CAD systems), so an additional category of non-nodules > 3mm was also adopted. Objects that would fall in this category would include scars, mucus plugs and other objects that are clearly not cancers. For these objects, again, a simple marking of the approximate centroid is provided.

To support these detailed and unambiguous annotations, the LIDC had to have two key technical developments. The first was the development of software tools that allowed the detailed markings of nodules and non-nodules including manual or semi-automated contouring tools. The

second was the development of a common file format that could contain detailed descriptions of nodule boundaries that was portable between the different marking tools. The former required some technical developments on the parts of several groups within the LIDC and resulted in software from three different groups that all accomplished the same ultimate task of marking nodules and non-nodules. The latter required the development of a common xml schema that allowed all three software tools to create a common file format that accurately and unambiguously described nodule contours and centroids as well as nodule annotations.

The LIDC process model of a detailed blinded and unblinded reading of a case has been implemented. The data collection process has been shown to be robust and practical and to provide the best estimate of nodule location and extent while retaining information about the variability between readers that exists. The technical developments and database design elements have been completed and implemented for the LIDC.

While the objectives of the RIDER database may be different from those of the LIDC, the technical infrastructure and data collection process of the LIDC can be easily adapted to that of the RIDER project. This will provide a mechanism to collect not only image data, but detailed and unambiguous information about radiologists' interpretation of the image data that can serve as input information for systems to assess changes in response to therapy. That is, radiologists can still identify lesions (nodules, masses or even other categories of desired objects to mark), can outline them or mark them in some other fashion (a single diameter such as RECIST marking or multiple diameters such as WHO or even the entire volume from which these other markings can be derived), and can annotate them in terms of characteristics (such as shape, margin, etc.). This can be done in a blinded, independent reading fashion or the process model can be extended to include the LIDC model of blinded and unblinded reads; the existing technical infrastructure can support either approach.

5.4.2 Validation Software Tools. The RIDER database leverages the NCI Center for Bioinformatics (NCICB) informatics infrastructure backbone, which includes vocabulary services and meta-data management tools.

One of the problems confronting the biomedical data management community is the panoply of ways that similar or identical concepts are described. Such inconsistency in data descriptors (meta-data) makes it nearly impossible to aggregate and manage even modest-sized data sets in order to be able to ask basic questions. The NCI, together with partners in the research community, developing common data elements (CDEs) that are used as meta-data descriptors for NCI-sponsored research and for the caCORE applications. The Data Standards Repository (caDSR) is a database and tool set that the NCI and its partners use to create, edit and deploy the CDEs.
(http://ncicb.nci.nih.gov/NCICB/infrastructure/cacore_overview/cadsr).

Management of RIDER CDEs in the caDSR will facilitate: 1) an available repository of information about the data in studies, 2) creation of questionnaires and forms, 3) comparison of data elements from different studies, 4) utilization of common data elements and thesauruses from other disciplines, and 5) ensuring that other disciplines can benefit from knowledge about RIDER domain.

RIDER elements that will be available through the RIDER database public interface have been entered and are accessible through the NCI CDE Browser found under the hierarchy:CIP/Classifications/I3Projects/RiderProject. (http://cdebrowser.nci.nih.gov/CDEBrowser/)

The LIDC and IDRI have also used these to great effect; one can see the LIDC CDEs can be found under the same web page hierarchy: CIP/Classifications/LIDC. Having a common vocabulary and common data elements reduces the confusion and frustration common in multi-center and multidisciplinary studies.

In addition to RIDER leveraging the NCICB informatics infrastructure, several other projects within the NCI are looking to use the RIDER database as a use case (such as the In Vivo Imaging Workspace chaired by Eliot Siegel an NCI IPA). We expect significant coordination and synergy between these projects which will result in the RIDER project being kept abreast of developments within caBIG and even being able to help direct those projects as they develop.

Patient Confidentiality is always an issue in these kinds of studies, but with the MIRC software tools for the field centers, sites are able to remove protected health information (PHI) fields easily while maintaining unique identifiers for both imaging studies and for other meta-data. This allows tracking of other necessary information (demographics, clinical results, treatment information) without compromising patient confidentiality and ultimately the linking of image data to treatment and outcome data.

All of these tools will allow investigators to focus their efforts on the key aspects of the project – what data elements to collect and how to collect them –without having to focus on the infrastructure necessary to support this project.

5.6 Standardized Methods for Software Performance

Design issues in standard measurements of software tool. In addition to the research issues as outlined in section 5 above, there are matters of the design of the process by which the performance of software tools are to be benchmarked and evaluated. These design issues are found in a variety of software performance applications, ranging from software assisted medical diagnosis to the verification of biometric software such as in face or finger print detection developed at NIST. The discussion of evaluation design reflects the results with biometrics reflected in the references at the end of this report.

Factors that need review in collaboration with NIST and FDA include the following:

- **Data Sequestration Requirements:** The users of lung cancer software diagnostic tools need confidence in the performance of the algorithms on the broad variety of imagery likely to be encountered in a clinical setting. The application of the tools to imagery distinct from that on which the tools were developed provides confidence that performance is not the result of tuning to a particular collection of cases.
- **Training and Evaluation Database Requirements:** In the design of performance benchmarking and evaluation, it is useful to separate the data on which the software operates into two sets, a training set and an evaluation set. These classes are distinguished in that prior to an evaluation the training set, including all annotations, is broadly available to the algorithm developers while the evaluation set is sequestered, and is available only to an independent evaluation team, and not to the developers. The process of sequestration is used to assure that in the evaluation indicates the performance on new future imagery rather than being "tuned" to a particular set of cases. The size of each data set will affect the statistical measures of confidence in the results.
- **Governorship and Maintenance of Data Archive:** The developers of software tools will undoubtedly want access to the sequestered imagery and associated annotations once an evaluation is complete. Providing access allows the developer to better understand performance and to improve the tools. It is likely that there will be a number of evaluations, implying a continuing need to collect new cases for inclusion in the sequestered archive, as cases migrate into the training (open) portion of the archive.

Some NIST URL's of interest related to measurement standards includes the following:

http://www.frvt.org/DLs/FERET7.pdf

http://www.biometricscatalog.org/documents/Phillips%20FRGC%20-%20Feb%202005-5.pdf

http://polymers.nist.gov/uploads/dunkers0104.pdf

5.7 Engagement and Scientific Role of Industry Stage Holders

5.7.1 Pharma: Potential Image Data Sources: Industry Drug Trials. The intent is to share the RIDER White Paper with industry stakeholders through the office of the FNIH. One of the goals is to explore access to privately supported drug trails and in particular trial data were the drug is already approved since the image and much of the meta-data has little intellectual property (IP). Often this will involve and CRO that manages the drug trial. This data access for say lung cancer drug trials would facilitate the following:

(a) Accelerate the data collection and completion time line of the RIDER project, and permit phased collection of image data where changes in the imaging technology can be accommodated.

(b) Permit the accrual of representative data both nationally and internationally.

NCI has therefore collaborated with the RSNA to provide open source software tools (MIRC) so that both image and in the near future metadata, to be de identified and collected through most fire walls, in an efficient manner with minimal over head. However, the following additional data may be needed and further discussions with pharma consortium are necessary as follows:

(a) Imaging Platform Quality Control Data (Q/C) collected at the clinical sites, where the rigor of the Q/C measurements needs to be reviewed in the context of the goals of the RIDER resource serving as a reference standard, in particular for PET-CT.

(b) Metadata collection, in terms of scope, methods for transfer and archiving.

(d) Outcome data as described in the next section, with an understanding of the longer time table and related IP and patient confidentiality issues.

5.7.2 Correlation of Change Analysis Data with Clinical Outcomes. The RIDER pilot project was originally designed to initiate the development of a database to optimize and compare the relative performance of change analysis tools, and to collect data from on going drug trials as cost effectively as possible, by inclusion of a range of lung cancer drug trials. The rationale for this limited scope was the view that we must first fully optimize and compare the performance of different of change analysis tools that use many computed features such as lesion volume, shape or texture or other pixel based features. This resource would be the first step in evaluation of tools for measuring change analysis, but does not answer how the results for the software tool that is selected correlates with clinical outcome for a specific drug. Thus there is a need to evaluate the selected methods in on going NCI or privately funded clinical trial. However the following issues need further discussion with the potential pharma stakeholders:

(a) Since the anticipated changes in CT or PET scans are expected to be specific to a given class of drugs as outlined in section 3.2, the possibility for sharing the results of change analysis data could be explored for a given drug class. Since the pharma consortium has agreed to explore sharing of software tools for change analysis, this process would help Pharma identify the best software tools to employ in their submissions to FDA for a targeted class of drugs.

(b) The potential role of the NCI funded RIDER academic partners, FDA and NIST in these efforts needs to be reviewed by Pharma, as there may be a need to establish a consensus on the statistical methods to be employed.

5.7.3 Device and Software Industry: Standards. The IDRI PPP developed by the FNIH was initiated in the fall of 2005. The steering committee structure has greatly facilitated scientific interactions between academia and the industry members that have included further improvements in the imaging protocol, design and content of the IDRI database and its functionality, in particular in the context of regulatory approval of application specific tools. http://www.fnih.org/partners/research_environment/IDRI.shtml

There is therefore a need to review the following standardization issues with industry:

(a) Level of interest in both creating standardized methods for image annotation and mark up as well as methods for evaluation of change analysis tools. The IDRI PPP does not support the development of standardized methods for assessment of CAD methods for lung cancer diagnosis and treatment. However since the end user of these tools involved pharma, the are particularly interested in standardized assessment of these software tools to permit an objective choice for drug trial submission to the FDA.

(b) Several of the software industry partners are particularity interested in development of data integration tools and integrated solutions for clinical decision making as stated in a recent BECON BISTI symposium 2004.
(http://www.becon.nih.gov/symposium2004.htm).
As reviewed in section 3.2, there is a need to implement data integration methods for lung cancer using both laboratory and imaging derived biomarkers, and drug response is an excellent clinical model to explore. This symposium and others informatics symposiums held in 2006, including the RSNA New Horizon's lecture, have stressed the need for standardized

methods for data collection and analysis to reduce the sources of uncertainty in implementation of data integration methods. NCI CIP and caBIG is exploring targeted workshops to engage the IT industry interested in this problem.
(c) The ACR is convening a meeting to explore uniform acquisition of image data across imaging systems as suggested by NCI. As outlined in section 6.3, NCI is exploring how to standardize data collection for drug trials for emerging imaging methods such as combined MRI/Optical. There is therefore an opportunity to explore further industry collaboration with RSNA IHE, NEMA, NIST and NIH, with a few of improving data collection for databases and future clinical drug trials.

6. NCI and NIH Complementary Initiatives to the RIDER Project.

6.1 PAR: Industry-Academic Partnerships. A new PAR entitled "Industry Academic Partnerships for Translational Research" is under review at NCI with a potential publication date by summer 2006, using an R01 funding mechanism. The primary focus of this planned PAR is to enhance translational research for targeted clinical investigations. One of the goals is the development of public resources to accelerate translation research and creation of standardized methodology for data acquisition and analysis. The PAR should provide a means to extend the RIDER project using investigator initiated grant submissions over the next several years that may include partnerships between academia, the device and pharmaceutical industries, and scientists from other government agencies (FDA, NIST, and NSF).

6.2 The ACR UPICT Initiative. The American College of Radiology (ACR) and NCI Cancer Imaging Program mutually recognize imaging's opportunity to contribute key tumor data in clinical trials. But imaging's reproducibility as quantitative data is often impaired by a lack of uniformity in protocols across the many sites required for patient accrual. Thus consensus on imaging protocols used in multi-site clinical trials is needed, particularly for the most widely used oncology tools such as CT, MRI, DC-MRI and FDG-PET. Hence the two organizations are proceeding in a complementary manner to achieve a goal of developing guidelines for imaging acquisition in clinical trials by a process entitled 'UPICT' – Uniform Protocols for Imaging in Clinical Trials. The ACR has convened a broad spectrum of participants that include diagnostic radiologists, radiation oncologists, medical physicists, clinical trial experts, government agency representatives and appropriate industry representatives. The bedrock of this effort recognizes that existing installed imaging technology is varied and must reasonably accommodate research imaging procedures in the context of usual care. To do so it must acknowledge rapid progress of technology and incorporate change management so as to provide periods of stability through predictable, pre-announced version control. A goal has been set to provide protocols that can be disseminated by late 2006. (http://upict.acr.org).

6.3 NCI CTEP: Reorganization of Drug Therapy Trials. In January 2004, the Director of the National Cancer Institute (NCI) established the Clinical Trials Working Group (CTWG) to advise the National Cancer Advisory Board (NCAB) on whether and in what ways the NCI-supported national clinical trials enterprise should be restructured to realize the promise of molecular medicine for advancing oncologic clinical practice in the 21st century. In June 2005 NCI issued its' Report of the Clinical Trials Working Group of the National Cancer Advisory Board, Restructuring the National Cancer Clinical Trials Enterprise, (http://integratedtrials.nci.nih.gov/ and http://integratedtrials.nci.nih.gov/ict/CTWG_report_June2005.pdf)

The CTWG was a broadly constituted panel with experts from academic research institutions, community oncology practices, the pharmaceutical and biotechnology industries, cancer patient advocacy groups, NCI, the Food and Drug Administration (FDA), and the Centers for Medicare and Medicaid Services (CMS). The CTWG first reached consensus on critical goals for designing a restructured national clinical trials enterprise that is not only more efficient and coordinated but founded on the best science. One of these goals was to improve standardization of tools and

procedures for trial design, data capture, data sharing, and administrative functions to minimize duplication of effort. Moreover, it stated that the evaluation of novel targeted therapies, designed to be effective against cancers with a specific molecular profile, depends on synergistic integration of treatment protocols with modern molecular diagnostic and imaging techniques. The RIDER initiative is poised to take on a highly integrated and supportive role in this effort.

6.4 NCI IRAT: Engagement of Cancer Center Imaging Programs.

NCI Image Response Assessment Team (IRAT) supplement awards have been administratively funded in the fall of 2005 to eight imaging teams in Cancer Centers to advance the role of imaging in assessment of response to therapy. In the 2005 NCI-sponsored RFA, 31 Cancer Center respondents provided cohesive plans to enhance involvement in quantitative analysis, interpretation, and integration of imaging data in response to therapy trials and eight of them were funded. Their plans also included a means for regular dissemination and communication of these methods with IRAT's at other institutions in a process managed by the American Association of Cancer Institutes (AACI). The long-term objective of these awards was to increase clinical collaboration between imaging scientists and oncologic investigators at Cancer Centers. The teams are identifying new oncologic imaging research opportunities in clinical trials that warrant multi-center clinical investigations and plan to integrate imaging data as potential biomarkers or candidate surrogate markers in clinical therapeutic trials. FY 2005 funding of eight administrative supplement awards totaled $2M. FY 2006 and FY 2007 funding will continue through the Cancer Center program. *NCI CIP program is thus exploring how to link the activities of the IRAT program with the RIDER project and the NCI Imaging Workspace.*

7. Authors

Academic: PI and Co PI's only: Alphabetical order

Denise Aberle MD., University of California Los Angeles
Sam Armato PhD., University of Chicago
Luc Biduat PhD., M. D. Anderson Cancer Center
Eric Hoffman PhD., University of Iowa
Ella Kazerooni MD., University of Michigan:
Heber Mc Mahon MD., University of Chicago
Geoffrey McLennan MD., University of Iowa: Chair of LIDC and RIDER
Mike McNitt-Gray PhD., University of California Los Angeles
Chuck Meyer PhD., University of Michigan: Co Chair of LIDC and RIDER
Reginald Munden M D., M D Anderson Cancer Center
Anthony Reeves PhD. Cornell University
Larry Schwartz MD., Memorial Sloan Kettering Cancer Center
David Yankelevitz MD., Cornell University
Binsheng Zhao PhD., Memorial Sloan Kettering

Academic: Support Staff

University of California Loss Angeles (UCLA)

Matthew S. Brown, PhD
Richard C. Pais, BS
David P-Y Qing, BS

Cornell:

Ali Farooqi, MD
Matthew Cham, MD
Daniel Max, BS
Alberto Biancardi, PhD

University of Iowa

Radiologists: Brian Mullan MD
Edwin van Beek MD PhD
Kim Sprenger RN
Christopher Piker PhD
Lisa Hudson
Jered Sieren

University of Michigan
Peyton H. Bland
Gary Laderach

NCI Staff
Larry Clarke PhD, CIP, Branch Chief: Science Officer
Barbara Croft PhD,CIP, Program Director
Carl Jaffe MD, CIP, Branch Chief
Gary Kelloff MD, CIP, Special Assistant (OBQI)
Lori Dodd PhD, CIP, Bio-statistics Branch
Dan Sullivan MD, CIP, Associate Director of CIP
caBIG IT Support
John Freymann BS
James Lou PhD
David Kupfersmid MS

NCI Visiting Scientist
Paul Kinahan PhD., University of Washington: PET-CT Consultant
Ron Gottlieb MD., Rose Park Hospital**:** Annotation Consultant
Eliot Siegel MD., University of Maryland: CaBIG, Imaging Workspace.

NIBIB
Belinda Seto PhD, Deputy Director

CRPF
Jim Mulshine MD
Bo Aldge

RSNA
Gary Becker, MD
John Perry (RSNA-MIRC Development)

Other Federal Agencies.
FDA
Kyle Myers, PhD
Nicholas Petrick, PhD
Bob Wagner, PhD

FDA RIDER Research Fellows
Marios Gavielides PhD: NCI NIBIB Fellow
Lisa Kinnard PhD: NIBIB Fellow

NIST
Information Technology
Charles Fenimore PhD
Martin Herman PhD
Imaging Hardware
Lisa Karam PhD

Brian Zimmerman, PhD
Paul Bergstrom, PhD
Charles Clark, PhD

8. References

Joint Publications: LIDC and RIDER

Dodd, L.E., et al., Assessment methodologies and statistical issues for computer-aided diagnosis of lung nodules in computed tomography: contemporary research topics relevant to the lung image database consortium. Acad Radiol, 2004. 11(4): p. 462-75.

Armato, S.G., 3rd, et al., Lung image database consortium: developing a resource for the medical imaging research community. Radiology, 2004. 232(3): p. 739-48.
Tran LN, Brown MS, Goldin JG, Yan X, Pais RC, McNitt-Gray MF, Gjertson D, Rogers SR, Aberle DR. Comparison of treatment response classifications between unidimensional, bidimensional, and volumetric measurements of metastatic lung lesions on chest computed tomography. Acad Radiol. 2004 Dec; 11(12):1355-60.

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

Armato SG III, McNitt-Gray MF, Reeves AP, Meyer CR, McLennan G, Clarke LP, et al. The Lung Image Database Consortium (LIDC): An evaluation of radiologist variability in the identification of lung nodules in CT scans. (in preparation, to be submitted to Radiology)

Related Juried Publications: Academic Sites since 2003

University of California Los Angles (UCLA)

Petkovska I, Shah SK, McNitt-Gray MF, Goldin JG, Brown MS, Kim HJ, Brown K, Aberle DR. Pulmonary nodule characterization: A comparison of conventional with quantitative and visual semi-quantitative analyses using contrast enhancement maps. Eur J Rad. 2006 Apr 6

Brown MS, Goldin JG, Rogers S, Kim HJ, Suh RD, McNitt-Gray MF, Shah SK, Truong D, Brown K, Sayre JW, Gjertson DW, Batra P, Aberle DR. Computer-aided lung nodule detection in CT: results of large-scale observer test. Acad Radiol. 2005 Jun;12(6):681-6.

Shah SK, McNitt-Gray MF, De Zoysa KR, Sayre JW, Kim HJ, Batra P, Behrashi A, Brown K, Greaser LE, Park JM, Roback DK, Wu C, Zaragoza E, Goldin JG, Suh RD, Brown MS, Aberle DR. Solitary pulmonary nodule diagnosis on CT: results of an observer study. Acad Radiol. 2005 Apr;12(4):496-501.

Brown MS, Shah SK, Pais RC, Lee YZ, McNitt-Gray MF, Goldin JG, Cardenas AF, Aberle DR. Database design and implementation for quantitative image analysis research. IEEE Trans Inf Technol Biomed. 2005 Mar;9(1):99-108.

Tran LN, Brown MS, Goldin JG, Yan X, Pais RC, McNitt-Gray MF, Gjertson D, Rogers SR, Aberle DR. Comparison of treatment response classifications between unidimensional, bidimensional, and volumetric measurements of metastatic lung lesions on chest computed tomography. Acad Radiol. 2004 Dec;11(12):1355-60.

Brown MS, Goldin JG, Suh RD, McNitt-Gray MF, Sayre JW, Aberle DR. Lung micronodules: automated method for detection at thin-section CT--initial experience. Radiology. 2003 Jan;226(1):256-62.

Brown MS, McNitt-Gray MF, Goldin JG, Suh RD, Sayre JW, Aberle DR. Patient-specific models for lung nodule detection and surveillance in CT images. IEEE Trans Med Imaging. 2001 Dec;20(12):1242-50.

University of Chicago
Armato SG III, Altman MB, La Rivière PJ: Automated detection of lung nodules in CT scans: Effect of image reconstruction algorithm. Medical Physics 30: 461-472, 2003.

Armato SG III, Altman MB, Wilkie J, Sone S, Li F, Doi K, Roy AS: Automated lung nodule classification following automated nodule detection on CT: A serial approach. Medical Physics 30: 1188-1197, 2003.

Armato SG III: Image annotation for conveying automated lung nodule detection results to radiologists. Academic Radiology 10: 1000-1007, 2003.

Suzuki K, Armato SG III, Li F, Sone S, Doi K: Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography. Medical Physics 30: 1602-1617, 2003.

Armato SG III, Oxnard GR, MacMahon H, Vogelzang NJ, Kindler HL, Kocherginsky M, Starkey A: Measurement of mesothelioma on thoracic CT scans: A comparison of manual and computer-assisted techniques. Medical Physics 31: 1105-1115, 2004.

Armato SG III, Sensakovic WF: Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis. Academic Radiology 11:1011-1021, 2004.

Armato SG III, Roy AS, MacMahon H, Li F, Doi K, Sone S, Altman MB: Evaluation of automated lung nodule detection on low-dose CT scans from a lung cancer screening program. Academic Radiology 12:337-346, 2005.

Armato SG III, Oxnard GR, Kocherginsky M, Vogelzang NJ, Kindler HL, MacMahon H: Evaluation of semi-automated measurements of mesothelioma tumor thickness on CT scans. Academic Radiology 12:1301-1309, 2005.

Agam G, Armato SG III, Wu C: Vessel tree reconstruction in thoracic CT scans with application to nodule detection. IEEE Transactions on Medical Imaging 24:486-499, 2005.

Roy AS, Armato SG III, Wilson A, Drukker K: Automated detection of lung nodules in CT scans: False-positive reduction with the radial gradient index. Medical Physics 33:1133-1140, 2006.

University of IOWA

Li, B., Christensen, G.E., Hoffman, E.A., McLennan, G., Reinhardt, J.M. Establishing a normative atlas of the human lung: intersubject warping and registration of volumetric CT images. Acad Radiol. 2003 Mar; 10 (3): 255-65. PMID: 12643552

Aykac D, Hoffman EA, McLennan G, Reinhardt JM. Segmentation and analysis of the human airway tree from three-dimensional X-ray CT images. IEEE Trans Med Imaging. 2003 Aug.;22(8):940-50. PMID: 12906248

Hoffman EA, Reinhardt JM, Sonka M, Simon BA, Guo J, Saba O, Chon D, Samrah S, Shikata H, Tschirren J, Palagyi K, Beck KC, McLennan G. Characterization of the interstitial lung diseases via density-based and texture-based analysis of computed tomography images of lung structure and function. Acad Radiol. 2003 Oct; 10(10):1104-18. PMID 14587629

Judson MD, Baughman RP, Thompson BW, Teirstein AS, Terrin ML, Rossman MD, Yeager H Jr., McLennan G, Bresnitz EA, DePalo L, Hunninghake G, Iannuzzi MC, Johns CJ, Moller DR, Newman LS, Rabin DL, Rose C, Rybicki BA, Weinerger SE, Knatterud GL, Cherniak R; ACCESS Research Group. Two year prognosis of sarcoidosis: the ACCESS experience. Sarcoidosis Vasc Diffuse Lung Dis. 2003 Oct; 20(3):204-11. PMID 14620163

Tawhai MH, Hunter P, Tschirren J, Reinhardt J, McLennan G, Hoffman EA. CT-based geometry analysis and finite element models of the human and ovine bronchial tree. J Appl Physiol. 2004 Dec; 97(6):2310-21. Epub 2004 Aug 20. PMID 15322064

Hoffman EA, Clough AV, Christensen GE, Lin CL, McLennan G, Reinhardt JM, Simon BA, Sonka M, Tawhai MH, van Beek EJ, Wang G. The comprehensive imaging-based analysis of the lung: a forum for team science. Acad. Radiol. 2004 Dec; 11(12):1370-80. PMID 15596375

Suter M, Reinhardt J, Montague P, Taft P, Lee J, Zabner J, McLennan G. Bronchoscopic imaging of pulmonary mucosal vasculature responses to inflammatory mediators. J Biomed Opt. 2005 May-Jun; 10(3):034013 PMID: 16229657.

University of Michigan

Brock, K, J Balter, L Dawson, M Kessler, and C Meyer (2003) Automated generation of a four-dimensional model of the liver using warping and mutual information, Medical Physics 30(6):1128-1133.

Meyer, CR, H Park, JM Balter, and PH Bland (2003) Method for quantifying volumetric lesion change in interval liver CT examinations, IEEE Transactions Medical Imaging 22(6):777-781.

Park, H, PH Bland, and CR Meyer (2003) Construction of an abdominal probabilistic atlas and its application in segmentation, IEEE Transactions on Medical Imaging 22(4):483-492.

Neemuchwala, H, AO Hero, PL Carson, and CR Meyer (2004) Local feature matching using entropic graphs, in Proceedings of Intl. Symp. Biomed. Imag. (ISBI), Arlington, Va.: 704-707.

Park, H, PH Bland, KK Brock, and CR Meyer (2004) Adaptive registration using local information measures., Medical Image Analysis 8(4):465-473.

Narayanan, R, JA Fessler, H Park, and CR Meyer (2005) Diffeomorphic nonlinear transformations: a local parametric approach for image registration, in Proceedings of Information Processing in Medical Imaging (IPMI), Lecture Notes in Computer Science 3565: 174-185.

Park, H, PH Bland, AO Hero, and C Meyer (2005) Least biased target selection in probabilistic atlas construction, in Proceedings of MICCAI'05, Lecture Notes in Computer Science 3750: 419-426.

Ma, B, R Narayanan, H Park, A Hero, P Bland, and C Meyer (2006) Comparing pairwise and simultaneous joint registrations of decorrelating interval exams using entropic graphs, submitted to MICCAI '06, Copenhagen, Denmark.

Meyer, C, B Moffat, K Kuszpit, P Bland, P Mckeever, T Johnson, T Chenevert, A Rehemtulla, and B Ross (2006) A methodology for registration of a histological slide and in vivo MRI volume based on optimizing mutual information, Molecular Imaging 5(1):16-23.

Related Publications: Cornell University:

A. P. Reeves, A. B. Chan, D. F. Yankelevitz, C. I. Henschke, B. Kressler, and W. J. Kostis. On Measuring the Change in Size of Pulmonary Nodules. IEEE Transactions on Medical Imaging, 25:435-450, April 2006.

W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke.
Three-Dimensional Segmentation and Growth-Rate Estimation of Small Pulmonary Nodules in Helical CT Images. IEEE Transactions on Medical Imaging, 22:1259-1274, October 2003.

C. I. Henschke, D. Shaham, D. F. Yankelevitz, A. Kramer, W. Kostis, A. P. Reeves, M. Vazquez, J. Koizumi, and O. S. Miettinen. CT Screening for Lung Cancer Significance of Diagnosis in its Baseline Cycle. Journal of Clinical Imaging, 30:11-15, January 2006.

W. Kostis D. F. Yankelevitz, A. P. Reeves, S. C. Fluture, and C. I. Henschke. Three-Dimensional Volumetric Measurement of Pulmonary Nodules. Radiology, 231(2):446-452, May 2004.

A. P. Reeves and B. M. Kressler. Computer-Aided Diagnostics. Thoracic Surgery Clinics, 14(1):243-244, February 2004.

Memorial Sloan Kettering Cancer and Research Center

Zhao B, Schwartz LH, Moskowitz CS, et al, Effect of CT Slice Thickness on Measurements of Pulmonary Metastases – Initial Experience. Radiology 2005; 234:934-939.

Zhao B, Schwartz LH, Lefkowitz RA, Wang L. Measuring Tumor Burden – Comparison of Automatic and Manual Techniques. Image Processing, Proc. SPIE 2004; 5370:1695-1700.

Zhao B, Schwartz LH, Moskowitz C, Ginsberg M, Rizvi AR, and Kris MG, Computerized quantification of tumor response in lung cancer – initial results. Radiology (in press)

Shoup M, Gonen M, D'Angelica MD, Jarnagin WR, DeMatteo RP, Schwartz LH, Tuorto S, Blumgart LH, Fong Y. Volumetric analysis predicts hepatic dysfunction in patients undergoing major liver resection. J of GI Surgery 2003; 7(3):325-330.

Dyke JP, Panicek DM, Healey JH, Meyers PA, Huvos AG, Schwartz LH, Tofts PS, GorlickR, Koutcher JA, Ballon D. Necrotic fraction estimation of primary bone tumors undergoing induction chemotherapy using dynamic enhanced MRI. Radiology. 2003; 228(1):271-278.

Schwartz LH, Mazumdar M, Wang L, Smith A, Marion S, Panicek DM, Motzer R. Response assessment classification in patients with advanced renal cell carcinoma treated on clinical trials. Cancer. 2003; 98(8):1611-9.

Schwartz LH, Mazumdar M, Brown W, Smith A, Panicek DM. Variability in Response Assessment in Solid Tumors: Effect of Number of Lesions Chosen for Measurement. Clin. Cancer Res. 2003; 9: 4318-4323.

Zhao B, Gamsu G, Ginsberg MS, Jiang L, Schwartz LH. Automated detection of small lung nodules on CT utilizing a local density maximum algorithm. J. Applied Clinical Medical Physics. 2003; 4(3):248-260.

Husband JE, Schwartz LH, Spencer J, Ollivier L, King DM, Johnson R, Reznek R. Evaluation of the response to treatment of solid tumors – A consensus statement of the International Cancer Imaging Society. Br J Cancer 2004; 90:2256-2260.

Mazumdar M, Smith A, Schwartz LH. A statistical simulation study finds discordance between WHO criteria and RECIST guideline. J Clin Epidemiol 2004;57:358-365

Yan J, Zhuang T, Zhao B, Schwartz LH. Lymph node segmentation from CT images using fast marching method. Comput Med Imaging and Graph 2004; 28(1):33-38.

Chung KY, Saltz B, Schwartz LH. Variability in response definition and confirmation practices in randomized phase II and III trials in colorectal cancer. J Clin Oncol 2004, ASCO Annual Meeting Proceedings (Post-Meeting Edition); 22(#14S):3594.

Zhao B, Schwartz LH, Moskowitz CS, Ginsberg MS, Wang L, Jiang L, Cooper C, Kalaigian J. Effect of CT slice thickness on measurements of pulmonary metastase-Initial experience. Radiology 2005; 234:934-939.

Mazumdar M, Smith A, Debroy PP, Schwartz, LH. A theoretical approach to choosing the minimum number of multiple tumors required for assessing treatment response. J Clin Epidemiol 2005; 58:150-153.

Morris MJ, Akhurst T, Larson SM, Ditullio M, Chu E, Siedlecki K, Verbel D, Heller G, Kelly WK, Slovin S, Schwartz L, Scher HI. Fluorodeoxyglucose positron emission tomography as an outcome measure for castrate metastatic prostate cancer treated with antimicrotubule chemotherapy. Clin Cancer Res 2005; 11(9):3210-3216.

Liu F, Zhao B, Kijewski PK, Wang L, Schwartz LH. Liver segmentation for CT images using GVF snake. Medical Physics 2005; 32: 3699-3706.

Zhao B, Schwartz LH, Moskowitz C, Ginsberg MS, Rizvi NA., Kris MG. Computerized quantification of tumor response in lung cancer – Initial results. Radiology (in press).

Yan J, Zhao B, Wang L, Zelenetz A, and Schwartz LH. Marker-controlled watershed for lymphoma segmentation in sequential CT images. Medical Physics (in press).

NCI CIP

Kelloff GJ, Krohn KA, Larson SM, Weissleder R, Mankoff DA, Hoffman JM, Link JM, Guyton KZ, Eckelman WC, Scher HI, O'Shaughnessy J, Cheson BD, Sigman CC, Tatum JL, Mills GQ, Sullivan DC, Woodcock J. "The progress and promise of molecular imaging probes in oncologic drug development." Clin Cancer Res. 2005 Nov 15;11(22):7967-85.

Jaffe CC, Measures of response: RECIST, WHO, and new alternatives. JCO Supplement in press, July 2006

Bezdek JC, Hall LO and Clarke LP. Invited Review: MR Image Segmentation Techniques Using Pattern Recognition. Medical Physics 8.Journal 1993, 20(4):1033-104

Clarke LP, Velthuizen RP, Camacho MA, Heine JJ, Vaidyanathan M, Hall LO, Thatcher RW and Silbiger ML. Review of MRI Segmentation: Methods and Applications. Journal of Magnetic Resonance Imaging 13(3):343-368, 1995.

Clarke LP, Velthuizen RP, Clark M, Gaviria G, Hall L., Goldgof D, Murtagh R, Phuphanich S and Brem S. MRI Measurement of Brain Tumor Response: Comparison of Visual Metric and Automatic Segmentation. Magnetic Resonance Imaging, 16(3), 271-279, 1998.

M. W. Vannier, E. V. Staab, L. C. Clarke, Medical image archives -- present and future, in: H. U. Lemke, M. W. Vannier, K. Inamura, A. G. Farman, J. H. C. Reiber (Eds.), Proceedings of the International Conference on Computer--Assited Radiology and Surgery (CARS 2002).

Croft et al: Committee for Review and Evaluation of the Medical Use Program of the Nuclear Regulatory Commission. Gottfried K-L D, Penn G, Eds. Radiation in Medicine: A Need for Regulatory Reform. Washington, DC: National Academy Press, 308 pp, 1996.

Houn F, Bright RA, Bushar HF, Croft BY, et al. Study design in the evaluation of breast cancer imaging technologies. Acad Radiol 2000; 7:684-692.

Eckelman WC, Tatum JL, Kurdziel KA, Croft BY. Quantitative analysis of tumor biochemisty using PET and SPECT. Nucl Med Biol. 2000 Oct;27(7):633-5.

NCI Visiting Scientists.

Kinahan PE, Hasegawa BH, and Beyer T. X-ray Based Attenuation Correction for PET/CT Scanners. Seminars in Nuclear Medicine vol. 33, pp 166-179,2003.

Alessio A, Kinahan PE, Cheng P, Vesselle H, and Karp JS, PET/CT Scanner Instrumentation, Challenges, and Solutions. Radiologic Clinics of North America, vol. 42, pp. 1017-1032, 2004.

Kinahan PE, Fessler JA, Alessio A, and Lewellen TK, Quantitative Attenuation Correction for PET/CT Using Iterative Reconstruction of Low-Dose Dual-Energy CT. In: 2004 IEEE Nuclear Science Symposium and Medical Imaging Conference, Rome, Italy, Oct 19-23, 2004, pp. 3285 - 3289, 2004.

Alessio A, Kinahan PE, and Lewellen TK, Improved Quantitation for PET/CT Image Reconstruction from System Modeling and Anatomical Priors with Clinically Feasible 3D Whole-Body PET Scanning. In: SPIE Medical Imaging 2005, San Diego, CA, Feb 5-9, 2005, pp. 695-703, 2005.

Siegel EL, Reiner BI. Filmless radiology at the Baltimore VA Medical Center: a 9 year retrospective. Comput Med Imaging Graph. 2003; 27 (2-3) : 101-9.

Siegel EL, Reiner BI. Introduction to the Paper by Eliot Siegel and Bruce Reiner, "Work Flow Redesign: The Key to Success When Using PACS" J Digit Imaging. 2003 Mar; 16 (1) : 163.

Siegel EL, Reiner BI. Work Flow Redesign: The Key to Success When Using PACS. J Digital Imaging 2003 Mar; 16 (1): 164-168.

Nagy P, Siegel E, Hanson T, Kreiner L, Johnson K, Reiner B. PACS reading room design. Semin Roentgenol. 2003 Jul;38(3):244-55

Siegel E. Ahead of the curve or out of the loop? Ten challenges to nuclear medicine interconnectivity.J Nucl Med. 2003 Oct;44(10):14N, 16N.

Reiner BI, Siegel EL, Siddiqui K. Evolution of the Digital Revolution: A Radiologist Perspective. J Digital Imaging 2004 Jan.

Andriole KP, Morin RL, Arenson RL, Carrino JA, Erickson BJ, Horii SC, Piriano DW, Reiner BI, Seibert JA, Siegel EL. Addressing the Coming Radiology Crisis – The Society for Computer Applications in Radiology Transforming the Radiological Interpretation Process (TRIP) Initiative. J Digital Imaging 2004 Dec;17(4) 235-243.

Reiner BI, Siegel EL, Hooper FJ, Siddiqui KM, Musk A, Walker L, Chacko. Multi-institutional analysis of computed and direct radiography: Part I. Technologist productivity. Radiology 2005 Aug:236(2):413-9.

Reiner BI, Salkever D, Siegel EL, Hooper FJ, Siddiqui KM, Musk A. Multi-institutional analysis of computed and direct radiography: Part II. Economic analysis. Radiology 2005 Aug:236(2):420-6.

Nagy P, Bowers G, Reiner BI, Siegel EL. Defining the PACS Profession: An Initial Survey of Skills, Training, and Capabilities for PACS Administrators. J Digital Imaging 2005 Oct 28

Reiner BI, Siegel EL, Siddiqui KM, Musk AE. Quality Assurance: The Missing Link. Radiology 2006 Jan ; 238(1):13-5.

FDA Scientists

S.V. Beiden, M.A. Maloof, R.F. Wagner. A General Model for Finite-Sample Effects in Training and Testing of Competing Classifiers. IEEE Trans. Pattern Analysis and Machine Intelligence (TPAMI) December 2003; 25 (12): 1561-1569.

W. A. Yousef, R. F. Wagner, and M. H. Loew, "Comparison of Non Parametric Methods for Assessing Classifier Performance in Terms of ROC Parameters," Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop, 2004; IEEE Computer Society, 2004.

W. A. Yousef, R. F. Wagner, and M. H. Loew, "Estimating the Uncertainty in the Estimated Mean Area Under the ROC Curve of a Classifier," Pattern Recognition Letters, vol. 26, pp. 2600-2610, 2005.

W. A. Yousef, R. F. Wagner, and M. H. Loew, "Assessing Classifiers From Two Independent Data Sets in Terms of The ROC Parameters: a Nonparametric Approach.," (in press) Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2006.

W. A. Yousef, R. F. Wagner, and M. H. Loew, "The Partial Area under the ROC Curve: Its Properties and Nonparametric Estimation for Assessing Classifier Performance," (in review) Pattern Recognition, 2005

S.V. Beiden, R.F. Wagner, G. Campbell. Components-of-variance models and multiple-bootstrap experiments: An alternative method for random-effects, receiver operating characteristic analysis. Acad Radiol 2000; 7: 341-349.

S.V. Beiden, R.F. Wagner, G. Campbell, C.E. Metz, Y. Jiang. Components-of-variance models for random-effects ROC analysis: The case of unequal variance structures across modalities. Acad Radiol 2001; 8: 605-615.

S.V. Beiden, R.F. Wagner, G. Campbell, H.P. Chan. Analysis of uncertainties in estimates of components of variance in multivariate ROC analysis. Acad Radiol 2001; 8; 616-622.

S.V. Beiden, R.F. Wagner, K. Doi, R. M. Nishikawa, M. Freedman, S-C B. Lo, X-W Xu. Independent versus sequential reading in ROC studies of computer-assist modalities: Analysis of components of variance. Acad Radiol 2002; 9: 1036-1043

R.F. Wagner, S.V. Beiden, G. Campbell, C.E. Metz, W.M. Sacks. Assessment of medical imaging and computer-assist systems: Lessons from recent experience. Academic Radiology 2002; 9: 1264-1277.

R.F. Wagner, S.V. Beiden, G. Campbell, C.E. Metz, and W.M. Sacks.Contemporary issues for experimental design in assessment of medical imaging and computer-assist systems. Proceedings of the SPIE Vol. 5034 (Medical Imaging 2003), p 213-224.

NIST Scientists

Albus, J.S., Quintero, R., Lumia, R., Herman, M., Kilmer, R. and Goodwin, K. "A Reference Model Architecture for ARTICS." Manufacturing Review, Vol. 4, No. 3, September 1991, 182-193.

John M. Irvine; Charles Fenimore; David Cannon; et al., "Development of a Motion Imagery Quality Metric," Proceedings of the American Society for Photogrammetry and Remote Sensing (ASPRS) Annual Meeting, 1-5 May 2006, Reno, Nevada, USA, May 2006

Charles Fenimore, John Irvine, David Cannon, et al., "Perceptual Study of the Impact of Varying Frame Rate on Motion Imagery Interpretability", SPIE Conference on Human Vision and Electronic Imaging XI, SPIE 6057-17, January 2006

C Fenimore, V Baroncini, T Oelbaum, TK Tan, "Subjective testing methodology in MPEG video verification", SPIE Conf Apps Digital Image Proc, Denver CO August, 2004.

C. Fenimore and A. Nikolaev, "Assessment of resolution and dynamic range for digital cinema", SPIE Proc on Image and Video Communications and Processing 2003, Santa Clara, CA, January 2003

C. Fenimore, "Mastering and Archiving Uncompressed Digital Video Test Materials" SMPTE Journal, v.110:10, pp726 – 735, New York, NY [2001]

Ravichandran, V. and Sriram, R.D., Toward Data Standards for Proteomics, Nature Biotechnology, Volume 23, Number 3, pages 373-376, March 2005.

Ravichandran, V., Vasquez, G.B., Srivastava, S., Verma, M., Petricoln, E., Lubell, J., Sriram, R.D., Barker, P.E., and Gilliland, G., Data Standards for Proteomics: Mitochondrial Two-Dimensional Polyacrylamide Gel Electrophoresis Data as a Model System, Mitochondrian, Volume 3, 327-336, 2004.

Allen, R. H. and Sriram, D., The Role of Standards in Innovation,Special Issue on "Innovation: The Key to Progress in Technology and Society," Journal Technological Forecasting and Social Change, 2000.

Sriram, R., Fenves, S., Subrahmanian, E., Rachuri, S., Product Information Exchange: Practices and Standards, Transactions of the ASME Journal of Computing and Information Science in Engineering, Vol. 5, No. 3, September 2005.

Bergstrom, P.M., "The Use of Atomic Data in Applications Involving Ionizing Radiation," in Atomic and Molecular Data and Their Applications, ed. by David R. Schultz, Predrag S. Krstic and Fay Ownby (Amer. Inst. Phys., Melville, NY, 2003) 636, p. 5.

Burns, D.T., O'Brien, M., Lamperti, P., and Boutillon, M., "Comparison of NIST and BIPM Medium-Energy X-Ray Air-Kerma Measurements," J. Res. Natl. Inst. Stand. Technol. 108, 383-389 (2003).

Mourtada, F.A., Soares, C.G., Seltzer, S.M., Bergstrom, Jr., P.M., Fernandez-Varea, J.M., Asenjo, J., and Lott, S.H., "Dosimetry Characterization for ^{32}P Source Wire Used for Intravascular Brachytherapy with Automated Stepping," Med. Phys. 30, 959 (2003).

Seltzer, S.M., and Bergstrom, Jr., P.M., "Changes in the U.S. Primary Standards for the Air Kerma from Gamma-Ray Beams," J. Res. Natl. Inst. Stand. Techno. 108, 359-381 (2003).

Seltzer, S.M., Lamperti, P.J., Loevinger, R., Mitch, M.G., Weaver, J.T., and Coursey, B.M., "New National Air-Kerma-Strength Standards for ^{125}I and ^{103}Pd Brachytherapy Seeds," J. Res. Natl. Inst. Stand. Technol. 108, 337-358 (2003).

Tesk A., and Karam, L. R., "NIST and Standards for Tissue Engineered Medical Products," in Tissue Engineered Medical Products (TEMPS), ASTM STP 1452, ed.: Schutte, E., Picciolo, G. L., and Kaplan, D. S., West Conshohocken, PA: ASTM International (2003).

Cessna, J.T., Golas, D.B., Unterweger, M.P., and Zimmerman, B.E., "Establishment of Transfer Standard for Holmium-166-DOTMP," Appl. Radiat. Isot., 60, 505-510 (2004)

Devic, S., Seuntjens, J., Hegyi, G., Podgorsak, E.B., Soares, C.G., Kirov, A.S., Ali, I., Williamson, J.F., and Elizondo, A., "Dosimetric Properties of Improved GAFchromic Films for Seven Different Digitizers," Med. Phys. 31, 2392-2401 (2004).

Järvinen, H., Cross, W.G., Soares, C., Vynckier, S., and Weaver, K., Dosimetry of Beta Rays and Low-Energy Photons for Brachytherapy with Sealed Sources, ICRU Report 72, J. of the ICRU 4, No. 2 (2004)

Kramer, G.H., Hauck, B.M., Marro, L., Inn, K.G.W., Unterweger, M., Hodge, P., "The Use of Autoradiography for Investigating the Distribution of Radioactivity in Lung Counter Calibration Sources," Health Phys. 86 (4), 359-364 (2004).

Mourtada, F.M., Soares, C.G., and Horton, J.H., "A Segmented ^{32}P Source Monte Carlo Model To Derive AAPM TG-43/60 Dosimetric Parameters For Intravascular Brachytherapy," Med. Phys. 31, 602-608 (2004).

Murphy, M.K., Piper, R.K., Greenwood, L.R., Mitch, M.G., Lamperti, P.J., Seltzer, S.M., Bales, M.J., and Phillips, M.H., "Evaluation of the New Cesium-131 Seed for Use in Low-Energy X-Ray Brachytherapy," Med. Phys. 31, 1529-1538 (2004).

Rao, D.V., Seltzer, S.M., and Bergstrom, P.M., "Compton Scattering Cross-Sections for Individual Subshells for a Few Elements of Biological Interest in the Energy Region 5 KeV – 10MeV," Radiat. Phys. Chem. 70, 479-489 (2004).

Rivard, M.J., Coursey, B.M., DeWerd, L.A., Hanson, W.F., Huq, M.S., Ibbott, G.S., Mitch, M.G., Nath, R., and Williamson, J.F., "Update of AAPM Task Group No. 43 Report: A Revised AAPM Protocol for Brachytherapy Dose Calculations," Med. Phys. 31, 633 (2004).

Zimmerman, B., Cessna, J., and Millican, M., "Applied Experimental Determination of Calibration Settings for Plastic Syringes Containing Solutions of Y-90 using Commercial Radionuclide Calibrators," Radiat. and Isot. 60, 511-517 (2004)

Cessna, J.T., Hammond, M.M., Schultz, M.K., Zimmerman, B.E., "Standardization of Astatine-211 by Liquid Scintillation Counting," Radiocarbon, (in press)

Chen-Mayer, H.H., and Tosh, R.E., "The NIST Room Temperature Water Calorimeter," Biomedizinsche Technik, 50, Supp., Vol 1, Part 2, 1370-1371 (2005)

Desrosiers, M.F., Fattibene, P., and Le, F., "An Absorbed Dose Map of Bone Tissue Treated with a Radiopharmaceutical in vivo," Health Phys. (in press)

Devic, S., Seuntjens, J., Sham, E., Podgorsak, E.B., Kirov, A.S., Schmidtlein, C.R., LoSasso, T., and Soares, C.G., "Precise Radiochromic Film Dosimetry Using a Flat-Bed Document Scanner," Med. Phys. 32, 2245-2253 (2005)

Kirov, A.S., Piao, J.Z., Mathur, N.K., Miller, T.R., Devic, S., Trichter, S., Zaider, M., Soares, C.G., and LoSasso, T., "The Three Dimensional Scintillator Dosimetry Method: Test for a ^{106}Ru Eye Plaque Applicator," Phys. Med. Biol. 50, 3063-3081 (2005)

Mo, L., Avci, D., Baldock, C., Cessna, J.T., James, D., Simpson, B., Van Wyngaardt, W.M., "Development of Activity Standard for ^{90}Y Microspheres," Appl. Radiat. Isot., 63, 193-199 (2005)

Niroomand-Rad, A., Chiu-Tsao, S.-T., Soares, C.G., Meigooni, A.S., and Kirov, A.S., "Comparison of Uniformity of Dose Response of Double Layer Radiochromic Films (MD-55-2) Measured at 5 Institutions," Physica Medica 21, 15-40 (2005)

Rivard, M.J., Butler, W.M., DeWerd, L.A., Huq, M. S., Ibbott, G.S., Melhus, C.S., Mitch, M.G., Nath, R., and Williamson, J.F., "Response to "Comment on Update of AAPM Task Group No. 43 Report: A Revised AAPM Protocol for Brachytherapy Dose Calculations," Med. Phys. 32, 1822 (2005)

Volkovitsky, P. and Gilliam, D.M., "Possible PET Isotope Production Using Linear Deuteron Accelerators", Nucl. Instrum. and Meth. in Phys. Res. A 548, 571 (2005)

Williamson, J.F., Butler, W., DeWerd, L.A., Huq, M. S., Ibbott, G.S., Li, Z., Mitch, M.G., Nath, R., Rivard, M.J., and Todor, D., "Recommendations of the American Association of Physicists in Medicine Regarding the Impact of Implementing the 2004 Task Group 43 Report on Dose Specification for Pd-103 and I-125 Interstitial Brachytherapy," Med. Phys. 32, 1424 (2005)