Workshop IX: Application of High Resolution CT Imaging Data to Lung Cancer Drug Development: Measuring Progress

May 3 - 4, 2012 – Hyattsville, Maryland • USA

2012 EXECUTIVE SUMMARY

Steering committee:

James L. Mulshine, Rush University Medical Center

Rick Avila, Kitware, Inc

David Yankelevitz, Mount Sinai School of Medicine

Thomas M. Baer, Stanford University

Raul San Jose Estepar, PhD, Brigham and Women's Hospital – Harvard Medical School

Carolyn R. Aldigé, Prevent Cancer Foundation

Convener: The Prevent Cancer Foundation

The Ninth Prevent Cancer Foundation Workshop was held May 3 and 4 in Hyattsville, MD to explore current opportunities to advance the integration of quantitative CT imaging for evaluation of early lung cancer. This precompetitive workshop series is a forum for cross disciplinary discussion and planning to develop strategies which allow for more rapid progress.

To open the meeting, Dr. James Mulshine of Rush University reviewed the current status of this field, outlining the four major rationales for this forum:

- Lung cancer, as the leading cause of cancer death, is a major public health challenge across the entire world.
- CT imaging of the lung is a rapidly improving medical imaging modality.
- A single, state-of-the-art CT imaging study garners vastly more data than can be manually processed for application to routine clinical management and so further development of software especially to assist in quantitative clinical assessments
- And finally, although there have been some incremental successes, drug development for lung cancer needs to be greatly improved both to better help patients and to allow continued investment from the pharmaceutical industry.

The published Phase II neoadjuvant, window-of-opportunity clinical trial led by Nasser Altorki using the drug **pazopanib** demonstrated how quantitative imaging can be integrated with molecular diagnostics to enhance information yield (1). This trial approach gives the pharmaceutical industry better mechanistic feedback about the results of exposure of untreated lung cancer to the experimental drug. This trial strategy of integrating before and after tissue sampling and primary tumor imaging evaluations in relation to drug exposure provides a critical opportunity to more fully define the behavior of the molecular target in the primary tumor tissue. If neoadjuvant window trials are done routinely with all new lung cancer drugs, systematic analysis of this aggregated information would accelerate our understanding of the molecular underpinning of lung cancer. This trial strategy would also aid in identifying appropriate companion diagnostics for specific targeted molecules to indicate the success of the drug exposure. This detailed response assessment on the effect of the targeted agents on untreated; early lung cancer cohort is a key direction in the development of science-based strategies for lung cancer therapy. From a drug development perspective, this approach may also shorten the time to market for a new drug. This may be particularly true if the applications are focused on early disease management, such as with adjuvant or neoadjuvant therapy, similar to what is happening in early breast cancer management.

According to the CDC, smoking accounts for $167 billion annually in health care expenditures, with a significant fraction of those costs related to lung cancer. With the aging of the US population, the total number of annual lung cancer cases is expected to rise by 50% over the next two decades. Yet there are also favorable dynamics such as with the rapid improvement of CT imaging of the lung. In addition there is more interest in development of software to assist in evaluating high resolution imaging studies of the lung even with brief duration of drug exposure. The image archives of such studies, along with growing collections of lung images linked to clinical outcome will facilitate the validation of robust quantitative imaging tools.

The challenge is that the lung is an extremely complex anatomic structure surrounded by other complex structures and early lung cancers emerge in a wide variety of shapes and textures. Nevertheless, quantitative CT imaging is being used in lung cancer drug trials (such as with the pazopanib experience) as well as in early lung cancer detection trials.

During the previous seven workshops, participants have contributed not only to these types of clinical trials activities, but to the development of lung imaging databases, publications around the developmental challenges in this field, and the development of open source quantitation tools. There has also been work on software evaluation using a challenge mechanism to advance the acceptance of quantitative imaging as a formal tool to apply in FDA regulatory settings. In this context, many of the

participants in this workshop also participate in the Radiological Society of North America's Quantitative Imaging Biomarker Alliance (QIBA). The QIBA pursues the broader vision of developing all major medical imaging platforms for quantitative applications, while the goal of the workshop has remained focused on CT imaging in lung cancer. A positive collaboration has emerged, since lung CT is one of the more rapidly evolving areas of quantitative imaging and the workshop brings a strong clinical application orientation to the effort (2).

A shared interest of Workshop participants is in advancing quantitative CT-based lung imaging towards defining the most efficient and meaningful application of this approach to objectively demonstrate the unique value of quantitative CT imaging to improve lung cancer clinical management. Most of the focus on validating quantitative CT in the lung has been on lung cancer therapeutics. However two recent developments provide an alternative to consider. In a Dutch-Belgian randomized trial of CT-based lung cancer screening (NELSON), the use of quantitative CT to evaluate for growth in a suspicious nodule was associated with accurate and efficient detection of clinically significant lung cancers (3). This finding becomes even more significant with the National Cancer Institute Director's recent announcement of the National Lung Screening Trial's (NLST) finding of a 20% mortality reduction benefit from CT-based lung cancer screening (4). There is no consensus on how to most efficiently conduct imaging with lung cancer screening; thus having a robust quantitative imaging measurement tool to find rapidly growing lung cancers would be a great benefit. How to objectively validate such an application is also an important topic for Workshop discussion.

Therefore there are two major applications that provide significant opportunities for quantitative imaging validation: one with drug development and the other with lung cancer screening diagnostic work up. In both of these areas, to develop, refine and then validate robust performance, the outstanding need is for a large number of clinical cases in which the CT image files are available and linked with long term clinical outcome. This outcomes link material allows the actual performance of candidate image measurement tools to be objectively defined. A critical issue for this Workshop is to consider the challenges entailed in advancing either of these two validations. A further opportunity is to consider how the image resources for one application can be cross applied to advance the validation of the other. For instance with the three largest screening trials, the NLST, NELSON and the International Early Lung Cancer Action Project (I-ELCAP), there are over 90,000 cases of serial CT images for which clinical outcome is known (3-5). This data forms a profoundly valuable public health resource as a validation matrix and could greatly accelerate the development of precise quantitative imaging measurement tools.

The integration of quantitative imaging tools is facilitated by the emergence of electronic medical records, allowing imaging information to be integrated with other sources of information about patient management. The Institute of Medicine has published a series of reports discussing how to use large datasets to allow for continuous process improvement, which it describes as the "Rapid Learning Healthcare System". This approach of aggregating large clinical datasets not only provides a strategic opportunity for further improving quantitative tools, but also focuses attention on the strategic challenge of holding back both quantitative imaging and "rapid learning". In advancing both these approaches, there is a critical need for the routine donation of clinical images and the associated clinical follow-up information. As was discussed with the Rapid Learning Healthcare System, there is a need for large databases to allow for the development of evidence-based clinical management and corresponding need for database aggregation and data mining tools (6). To support this, there is a need to communicate the importance of donation of clinical data and to cultivate in the public an awareness of the need for routine data donation. Educating the public as to the importance of clinical data and imaging donation to advance research is an important challenge. Progress in this effort is essential to enabling the rapid development of quantitative imaging. Considering how to communicate this challenge to the public is an important issue to consider in the Workshop Breakout sessions.

In summary, there are favorable recent developments for advancing progress with quantitative imaging. We will review for this Workshop both computational imaging and targeted drug development. We will also work to refine pre-competitive strategies to integrate computational imaging in drug development for lung cancer or for the related-tobacco induced lung disease, COPD. We will again visit issues related to the more rapid development of software tools to allow accurate serial measurement of lung nodules in large part by supporting the creation of more effective image/data archives. We will also consider strategies to accelerate progress and discuss shared approaches with the advocacy community to communicate the benefit of quantitative imaging.

Overview of Presentations

Dr. Marietta Scott of AstraZeneca UK presented next on the application of quantitative imaging in a well-documented database of lung cancer cases. In particular the study involved multiple readers who performed multiple measurements of the cancers at different times using multiple methods. These included 1D, 2D, 3D manual and 3D semi-automated. From their preliminary experience, they found that volumetric measures reduce within-reader coefficient of variance when compared like-for-like with 1D measures. Computer assisted mark-up reduces both within- and between-reader coefficient of variance in volumetric measures and they found no evidence of bias in repeated mark-ups. They did find that CT slice thickness does introduce bias that seems to be dependent on measurement

method. For example, there was a larger average difference in measures between different slice thicknesses for smaller nodules, due to slice availability, which makes intuitive sense.

Rick Avila from Kitware presented on a new approach for reducing CT measurement bias and variability using CT pocket phantoms and automated analysis software. A small pocket-sized CT calibration phantom was described that was designed to be simultaneously CT scanned with patients, thereby providing image quality metrics and monitoring with every CT study. This initial CT pocket phantom was successfully deployed in the Roche ABIGAIL study, a phase II lung cancer clinical trial evaluating the effectiveness of two bevacizumab dosing options. The ABIGAIL study data is currently being analyzed and is significant to lung cancer imaging research in that it will provide the first dataset documenting high resolution CT imaging variability in a multi-institution lung cancer clinical trial. In addition to the ABIGAIL study data, several experiments and studies with CT pocket phantoms were described that have demonstrated higher levels of CT imaging variability than previously acknowledged.

One important issue raised was the lack of an accepted definition for CT slice thickness, a key driver in quantitative imaging performance for lung cancer and many other diseases and conditions. Without a precise definition of this important imaging parameter, there will remain the potential for high variability in CT imaging performance. Several recent developments with pocket phantom technology were described, including the development of a smaller and easier to measure 2nd generation pocket phantom and the addition of dose measurement devices on the pocket phantom to allow for dose and image quality optimization and monitoring. The CT pocket phantom presentation concluded with a summary of the many opportunities that this new technology provides, including the ability to more precisely control CT imaging variability, optimize image quality and radiation dose, and improve volumetric measurements.

The next presentation reviewed issues with formal regulatory qualification of volumetric CT as an imaging biomarker. Dr. Enrique Aviles of the Critical Path Institute reviewed the status of the process to approach the FDA for this designation. The experience with the successful submission of a set of urinary kidney biomarkers for autosomal dominant polycystic kidney disease (ADPKD) was reviewed. It was clear from this example that a large effect size has a significant positive impact on the FDA biomarker qualification process and that CT lung biomarkers should similarly strive for submission of biomarker data indicating a large effect.
In a complementary presentation, Andrew Buckler of the QIBA presented the progress of the Volume CT group in moving toward qualification for CT volumetrics for lung cancer. He outlined that the QIBA mission is to improve the value and practicality of quantitative imaging biomarkers by reducing variability across devices, patients, and time. The transformation for imaging is the notion that

imaging involves *"measuring devices"* rather than *"imaging devices."* Data from QIBA supports the notion that the integration of volumetric imaging data could both reduce the time to achieve a meaningful clinical endpoint in a clinical trial or reduce the number of subjects required for accrual to an informative clinical trial. Either endpoint is of value to the drug development community. A number of strategic partnerships are collaborating with QIBA towards the qualification goal and this effort involves multiple components, from defining the imaging process to obtain consistent high quality images to the details of how the image quantitation should best be conducted.

Dr. David Yankelevitz of Mount Sinai Medical Center next outlined a number of challenges associated with accurately measuring tumor volumes. These included difference in scanners or in software as well as challenges related to the tumor itself. In regard to the scanners, there are inherent properties such as modulation transfer function (MTF) and point spread functions that differ, and there are also adjustable scan parameters such as slice thickness, pitch, field of view, and dose. All of these potentially can have an effect on volume estimation. In regard to differences in software, each package uses a particular approach to defining the edge of a tumor, and this decision is different between the various manufacturers; thus, depending on the software used, the boundary chosen can vary and this leads to differences in measured volume as well. In regard to the tumor itself, he pointed out the extent of volumetric change is greatly influenced by the initial size of the nodule in regard to the relationship between surface area and overall tumor volume. This occurs because the proportion of surface area relative to overall volume is greater for smaller nodules and therefore minor differences in choosing the boundary have a greater effect on the volume estimates of smaller nodules compared to larger nodules. For sub-centimeter nodules differences of even a single pixel along the entire surface can change volume assessments on the order of 100%. Also, the degree to which an edge can actually be defined is somewhat limited by the way the tumor interfaced with the non-tumor surrounding parenchyma. Thus, there is an inherent uncertainty associated with defining a particular edge.

Finally, with complex lesions comprised of both solid and infiltrative components, there are difficulties in accurately segmenting the extent of the primary tumor due to the extensive invasion of infiltrating strands of tumor into surrounding parenchymal structures where the actual tumor boundary cannot be defined using traditional radiological boundary measures based on either density or morphology. This reality may make volumetric analysis unrealistic for this class of lesions without the development of entirely new volumetric techniques. In the interim, standard measurement techniques such as RECIST continue to have an important role, particularly in this latter class of lesions.

Dr. Terry Yoo of the National Library of Medicine reviewed the historical challenges with developing shared research resources and highlighted the strategic need for large validating data sources. Dr. Yoo reviewed the NLM's experience not only with databases and open source reference measurement tools but with the culture and the scientific impact of open research strategies. Dr. Susan Wood of VIDA Diagnostics then reviewed the issues with improving the climate for the development of image processing tools from a small business perspective. Certain regulatory features contribute to Europe's having a more favorable climate for aspects of image processing research. Moreover, development efforts directed at therapeutic applications may allow for more favorable reimbursement dynamics, but the key early adoptors in this strategy may be clinicians and not radiologists. With this new approach, there is the requirement to understand the clinician's workflow and to ensure that the new tool is appropriate to maintain efficient clinical management. In this challenging setting, it is paramount to have adaptive collaborators to ensure all energy is flowing towards a successful product development rather than dissipated with infighting.

Dr. Raúl San José Estépar of Brigham and Women's Hospital presented an overview of computational imaging research for COPD. He reviewed the critical components of the COPD image pipeline analysis and illustrated the challenges using the initial results from the analysis of 6,000 subjects in the COPDGene cohort. The failure rate for the definition of the lung region is below 1% for inspiratory and expiratory scans, highlighting the robustness of current approaches. Challenges remain for the extraction of the airway lumen with 25% of cases for which airway segments smaller than third generation could not be obtained. Measurements of the bronchial thickness still remains a challenge for thickness below 1.2 mm, mostly limited by the scanner resolution and the variability of the point spread function across the scanning field.

Dr. Estepar led a panel that discussed the practical challenges with software development confronting the computer science community in automating lung image analysis. Wes Turner from Kitware presented the latest results that can be achieved by the open source package, the Lesion Sizing Toolkit (LST). LST implements a multi-feature nodule detection algorithm that behaves within a 4% volume measurement error for synthetic lesion. Clinical results for the LST were shown from a range of lesions. Complex lesion shapes, especially those closed to other thoracic structures were hard to quantitate. The composite of confounding shapes variability in the CT scanning protocol or in slice thickness are among the critical factors that increase the measurement variance in volumetric detection algorithms.

Finally, Anthony Reeves of Cornell University stressed that standard volumetric methods have been designed for small nodules. Large complex lesions shapes are hard to capture by just volumetric changes. He suggested the used of density-based metrics over a fixed size region as a surrogate of

volume changes. Slice thickness is another main confounding factor that has to be corrected for, such as volume is insensitive to displacements of the nodule in the out-of-plane axis. Very small nodules also require attention with detailed algorithms that correct for differences in segmentation. With in-plane capture, progress has been made in improving the reliability of the computational approaches that can lead to the adoption of standard profiles based on those approaches.

Dr. Larry Clarke of NCI next outlined the progress of the NCI sponsored Quantitative Imaging Network. This is a growing extramurally funded research mechanism designed to promote research and development of quantitative imaging methods for the measurement of tumor response to therapies in clinical trial settings, with the overall goal of facilitating clinical decision making. A growing number of centers of imaging excellence have been selected through the NIH peer review process to address common issues including data collection, data evaluation and informatics. A key goal is to pool resources in these areas so that sites can leverage their resources and prevent "siloing", a common problem in many multi-site initiatives. Dr. Clarke noted that he has attended a number of these PCF workshops in the past and outlined the support of an early PCF workshop in advocating to the NCI leadership for the development of the Reference Imaging Database to Evaluate Response (RIDER) mechanism that he successfully led.

Dr. Raúl San José Estépar introduced the progress milestones that have been achieved in the COPD imaging community as part of the definition of new markers that may lead to a better group stratification based on risk profiles. The qualification of the current methodology that is being applied revolves around the agreement between the different approaches and the definition of clinical association with the new CT-based biomarkers that are being developed. In regard to emphysema quantification, new approaches that use local density can enhance our understanding of disease by dissecting the involvement of different emphysema patterns. In particular, the quantification of mild centrilobular disease can be paramount to the assessment of early disease that is precluded by current densitometry based on global histogram. In regard to airway disease quantification, current approaches show moderate agreement with a limited association to other clinical covariates. This result suggests that more refined airway phenotypes have to be sought in order to assess airway disease from CT scans. There are also new analyses of pulmonary vascular remodeling based on blood volume distribution across vessel size that can shed light on the understanding of the pathophysiological involvement of vascular disease in patients with COPD, illustrating the wealth of information that can be extracted from CT scans. All the efforts in quantification of disease based on CT scans are driven by the need to control variability and the validation of the biomarkers that will be discovered.

Dr. Jered Sieren from the University of Iowa covered the first issue related to the need of controlling variability by means of phantoms. He presented the experience of his group in the design and implementation of the quality control phantom for multicenter studies like COPDGene and Spiromics. Control phantoms that are routinely scanned can provide reference values for the variations in density that have direct impact in densitrometric measurements, like emphysema quantification. For example, multi-center studies using quantitative CT to evaluate emphysema have observed that tracheal air CT numbers are significantly greater than the standardized -1000 Hounsfield Unit (HU) value for air. New phantoms that reproduce those issues are being developed as reflected by the modified COPDGene and Catphan 500 phantoms. Phantoms can also flag site-specific variations in scanner behavior due to malfunction or unpredicted protocol modifications. Besides the need of these types of phantoms to control known variability, some measurement variability has been encountered with commercial CT scanners that is not fully understood or characterized, stressing the need for further imaging methodology research.

Dr. Jessica Sieren of the University of Iowa reviewed the challenges with relating the macro of the imaging data with the micro of the higher resolution of the anatomical imaging data from pathological tissue evaluation. This included demonstrations of mapping the microscopic features to the relevant region within the clinical image using a defined post tissue processing protocol, which allows for quantitative histological measurements that potentially can relate back to the clinical image. This approach allows for rich three dimensional composite analysis of the precise nature of the disease process. This approach can potentially provide a bridge from *in vivo* animal studies of mechanisms and guide imaging efforts in human Phase I/II trials.

Dr. Claudia Henschke of Mount Sinai Medical Center then reviewed the lung injury data that could be extracted from the extensive work of I-ELCAP with its large, serially followed tobacco-exposed population. In that database, visual assessments of extent of emphysema as well as coronary artery calcification are available. Overall the risk of developing lung cancer increases with risk of emphysema. Similarly, inter-relationships were explored regarding the correlation of coronary artery disease and emphysema. This inter-relationship of all three illnesses has long been postulated to have a common underlying mechanism related to inflammation. In addition, as newer software tools are developed, they can be applied to this large database to further refine the measurements of emphysema and coronary artery disease. This speaks to the advantage of having a database that continues to grow and can keep up with changes in technology. As newer techniques develop and new findings are made, they can be applied to the database to see if they can more accurately predict the likelihood of developing cancer. This is most evident for emphysema where techniques have continuously evolved. Initially, the I-ELCAP database relied on radiologists' estimation of extent of

emphysema; now, a variety of computer generated measurements can be applied, which allows for better quantification of extent of emphysema as well as better methods to predict cancer.

Dr. Anthony Reeves of Cornell University opened the second day of the Workshop with an overview of the status of computer-assisted image processing approaches. He reviewed the topics of malignancy status based on morphologic features, growth rate assessment, and detection. In regard to characterization of tumors based on morphologic features, this has been found to be relatively unsuccessful both for 2D as well as 3D feature analysis. Growth rate assessment is very precise when done with care and with due regard for measurement error. Detection techniques are also improving, especially for lesions greater than 4 mm. For further development and with the hope for dramatic improvements for detection and growth measurement, the main barrier is that the databases for development remain small and not well characterized. Large, well annotated databases would have a dramatic impact on the field.

Dr. Nicholas Petrick of the US Food and Drug Administration then reviewed his research efforts in defining the sources of error in the process of course of quantitative CT imaging for both clinical trials and in routine clinical practice. Much of the data to be discussed comes from a collaborative effort, QIBA volume CT group, where the focus was to estimate the bias/variance of radiologists' estimates of the size of synthetic nodules for CT scans of anthropomorphic phantoms. This involved analysis of five different-shaped nodules with two different defined nodule densities (-10HU or +100HU). Nodules were placed within the synthetic vasculature of the phantom and evaluated across a range of acquisition parameters and analyzed by CT on two separate occasions. The data was then reviewed by a panel of six experienced radiologists using 1D, 2D, and 3D techniques. The analysis suggests that the relative bias is 15% for 1D, 18% for 2D and -1% for 3D. The results of the bias measurements were significantly different across the three techniques. This difference was also consistent with different slice thicknesses, but 3D volume at 0.8mm slice showed lower bias and variance compared to 5mm slice. This kind of rigorous validation approach is now being extended to issues of robustness of reader variability overall and with nodule segmentation.

Next, Dr. Jorge Gomez of the National Cancer Institute reviewed the development of the US-Latin America Cancer Research Network (US-LA CRN). This is a consortium of national cancer research groups across the US and Latin America to work jointly on clinical trials in major cancer sites that are a shared public health problem. Dr Gomez reviewed the progress in implementing a state-of-the-art clinical trials infrastructure which initially has been working on breast cancer trials but is now considering evolving to other organ sites such as lung. Lung cancer is a leading cause of cancer mortality in Latin America and there is enthusiasm for cooperating on trials with quantitative imaging using shared protocols for this cancer.

The next presentation, by Dr. Natasha Leighl of the Princess Margaret Cancer Center was a review of the targeted therapy research with neoadjuvant window-of-opportunity studies using quantitative imaging. Dr. Leighl outlined how in this Phase II setting with previously untreated patients, there is an opportunity to evaluate the impact of exposure of a targeted therapeutic and obtain both molecular and quantitative imaging feedback on the effect of the drug exposure. This approach provides an opportunity to better understand the actual mechanism of action of a targeted therapy as well as to establish potential mechanisms of drug resistance.

In the next presentation, Dr. David Mozley of Merck Imaging reviewed the early results with the integration of quantitative imaging into clinical trials workflow as one approach to improving the economics of drug development. The assumption in this work is that quantitative imaging could allow for a greater percentage of evaluable clinical trial participants, so that fewer patients may be needed for a trial or perhaps shorter trial duration would be required. In a rigorous evaluation of actual clinical trials data, measuring whole tumor volume allowed for automatic measurement of longest diameters, which could improve both RECIST and WHO determinations. Deploying server at worksite involves more complex analysis but could reduce central versus site evaluation discordance. This may result in less participant censoring and improved care and faster trial completion. While further research is required, early results are promising relative to improved workflows with regulatory-directed clinical trials.

Two breakout groups were formed to address both the technical and the organizational issues and opportunities with advancing and deploying tools for early lung cancer management.
Breakout Group A was titled "**Accelerating Qualification Through Process Optimization**" and was facilitated by Ricardo Avila, Andrew Buckler and Marietta Scott. This group of Workshop participants explored the general areas of open image archives, image analysis algorithms, software resources and the importance of lung cancer imaging research as a funding priority at NIH.

As has been emphasized at all seven previous lung cancer workshops, it was recognized that the need for large, high quality and open image archives with associated metadata is both critical and growing. One of the solutions to this is to approach the pharmaceutical industry to make openly available some of the many high quality datasets collected during clinical trials. This is a challenging request due to the cost that will be incurred by the pharmaceutical industry and the difficulty in defining a crisp, short-term, financial benefit. For this to have immediate impact on lung cancer researchers, we will need the pharmaceutical industry to explore providing datasets from retrospectively performed clinical trials. However, informed consent wording across the pharmaceutical industry has been varied and in many cases has not specifically addressed making data openly available. Obtaining new

informed consent signatures from a large number of past lung cancer patients, allowing open image archive use of their data, would be cost- and time-prohibitive. To help determine if it is possible to utilize some of these pharmaceutical study collections with their existing informed consent language, it was determined that we should consult with commissioned ethicists and allow them to thoroughly explore this issue. It was also recommended that we consult with the HIPAA privacy group in the Department of Health and Human Services to obtain their perspective on this difficult area. To help address this issue going forward, it was recommended that standard informed consent language permitting public research use of data is distributed to all groups considering conducting a clinical trial, including industry, government, and academia. Breakout group A also recommended that we work with patient advocacy groups to encourage patients to participate in lung cancer clinical trials that allow for the data to be used in open image archive collections.

Breakout group A also considered the use of "data papers" to encourage researchers to make datasets publicly available. A data paper provides a data archive and a description of the data archive to readers, including the methods of data collection and the significance of the collection. The Optical Society of America's Interactive Science Publishing (ISP) system already has the necessary infrastructure to support this type of publication (7). It was determined that we would like to encourage the use of data papers and several steps were outlined to achieve this. First, the group recommended that meetings are established with the NLM and major lung cancer imaging publishers, such as Radiology, Medical Physics, and more broadly Elsevier, to encourage the use of "data papers".

Another important open image archive area discussed by Breakout Group A was the "Give a Scan Project" run by the Lung Cancer Alliance. The Lung Cancer Alliance has successfully conducted a pilot project exploring the potential for patients to directly donate their datasets to an open archive in order to help accelerate lung cancer research. With the recent successful completion of the National Lung Cancer Screening Trial, it was recommended that participants in upcoming lung cancer screening efforts are afforded the opportunity to donate their data to "Give a Scan". It was also recognized that in order to handle the data curation and processing issues associated with image archive operations, identifying additional sources of funding is critical to the future success of the project.

Breakout group A also addressed the general area of image analysis algorithms. It was recommended that the lung cancer research community continue to explore novel approaches to improving the assessment of therapy response, such as spectral CT, CT perfusion, and CT calibration. However, retaining an emphasis on technologies that leverage the existing capabilities of deployed CT scanners will be critical to translation into clinical trials and practice and is therefore recommended.

For the software resource topic area, the group reviewed the need to more closely collaborate on software tools and models between lung cancer and COPD research efforts. An open source lesion sizing architecture and reference lung cancer sizing algorithm has been publicly available and actively disseminated for over two years. The National Alliance for Medical Image Computation (NA-MIC) project has disseminated open source code for the analysis of COPD for several years as well. It was recognized that collaboration and integration of these and other software resources have the potential to accelerate algorithm development in both fields. For a second consecutive year the breakout sessions identified the need to build a common software platform for lung cancer and COPD imaging research. To achieve this it was further recommended to discuss this topic with the National Library of Medicine as a future addition and direction for the Insight Segmentation and Registration Toolkit (ITK). It was also recommended that QIBA explore this area, given that there are profiling efforts underway for both lung cancer and COPD quantitative imaging.

A final area considered by the breakout group was the need for NIH leadership in thoracic imaging research. Several circumstances provided motivation for this topic area. CT scanner manufacturers continue to make rapid technological improvements in fundamental imaging characteristics, spectral acquisition, and more efficient dose utilization. In addition, the significance of high resolution CT chest imaging and advanced software technology continues to grow as CT screening studies, COPD studies, and cardiac studies demonstrate clear and unmatched clinical benefit. However, government funding of novel CT algorithms and applications for chest imaging is sparse and distributed across many NIH institutes and other government organizations. In addition, the output of these efforts is generally fragmented and disjointed, with insufficient collaboration at both the clinical and the software engineering levels. It was recommended that NIH, and in particular NCI and NHLBI, consider creating a targeted RFP for improving quantitative thoracic algorithms based on high resolution CT. It was also recommended to discuss the potential for an open source initiative on quantitative CT with the larger pharmaceutical industry. As with previous lung cancer workshops, the recommendations from this workshop will be translated into specific action items which will be worked on throughout the year.

The second breakout group was led by Dr. Tom Baer of the Photonics Institute, Stanford University, Carolyn Aldigé, Prevent Cancer Foundation and Laurie Fenton, Lung Cancer Alliance. The breakout was entitled, **Strategic Planning: Exploring Models for Implementing a National Infrastructure for Quantitative Imaging in Early Lung Cancer Management.** The goal of this group was to propose a procedure for implementing a national infrastructure for deploying CT based quantitative imaging technology for early lung cancer detection and management. The group leaders outlined a strategic planning process which included defining goals in developing this national infrastructure. The elements of this task included characterizing the requirements for nationwide, high throughput, low

cost, CT lung cancer screening facilities for early detection of lung cancer in high risk individuals. Another task was to explore potential cost efficiencies by combining the screening program with a smoking cessation and life style counseling program. The final task was to provide an integrated patient management process for positively screened patients involving best practices in regard to safety and efficiency.

As summarized in Figure 1, the Breakout group discussed who would be potentially interested in participating in such a developmental effort. The spectrum of stakeholders spans those that receive the service, those that deliver the service, those that pay for the service and finally those that regulate the delivery of the service.

Virtually all of the stakeholders share an interest in access to aggregated clinical data. This resource could be used in a global ongoing fashion to ensure continuous process improvement. The information could be used specifically to enhance our understanding of the cost effectiveness of the service, which ultimately impacts the reimbursement policies. These types of issues would be of particular interest not only to patients but for many of the professional societies for the many types of healthcare providers engaged in providing this kind of care. More textured clinical/imaging data will be essential for the lung cancer screening research and development professionals as well as to the relevant regulatory professional. Therefore it is evident that broad consortiums of stakeholders need to have mechanisms for communication and shared planning to allow this new public health service to evolve in a robust, economical and effective way.

As a tool to accelerate progress towards a national lung cancer screening program as a viable public health infrastructure, we propose an initial action plan. The elements of this plan include developing a program to demonstrate that the protocols employed in recently published screening trials can be cost-effectively translated into a public health setting. We contend this requires integration of the lung cancer screening process with smoking cessation advice and life style counseling programs. Further, from a quality and cost perspective it would be desirable to integrate with the screening centers a managed care program that will provide services to patients who screen positive for lung cancer. We outline a series of clinical settings where pilot demonstration projects can be conducted to allow a better picture of implementation feasibility across a cross section of patient populations. These populations could be served at sites that might include the Veterans Administration, a managed health care facility like Kaiser Permanente and a hospital serving a representative normally underserved low income population.
The demonstration trials would need to have clearly measureable goals. One would expect to provide a level of care comparable to the efficacy reported in the NLST trial in the three types of medical venue. Additional endpoints of the demonstration project would be to measure the observed impact

of smoking cessation programs when combined with screening. Finally, it would be critical to evaluate a cost/benefit model for operation of a high throughput screening center in the three medical venues.

In conclusion, two compelling clinical applications existing for integrating quantitative CT imaging into lung cancer research. In regard to the use of quantitative imaging for drug development, there are growing numbers of serial CT image/clinical outcome cases that are available as archival resources over the web. However, the number of useful, high resolution cases is still modest. This limitation is greatly constraining progress in the field, especially around the development of robust software tools. Renewed efforts are needed to encourage great imaging case donations, so that the aggregate size of this resource grows quickly.

The second application is the use of quantitative CT imaging in the diagnostic evaluation of suspicious cases of screen-detected lung cancer. This is an example using volume growth criteria as measure by CT image process serving as a surrogate of clinically aggressive cancer. This approach has been recently published in a report on diagnostic work-ups within the ongoing NELSON randomized screening trial. Interval nodule growth as initially reported by Yankelevitz and co-workers (8) is an important filter to minimize over diagnosis bias associated with CT-based lung cancer screening.
There is an opportunity in developing these applications in parallel, since the CT imaging cases from either the drug development or the screening case detection can be aggregated and used to cross validate software for either of the two target applications. Similarly, the use of imaging phantoms with quantitative CT provides a crucial opportunity to determine sources of measurement variance associated with CT imaging. Further dialogue is essential across the field, to get buy-in with the most informative phantom construct, so again wide use of the most informative phantom can accelerate systematic investigation of the sources of variance occurring with CT imaging for lung cancer clinical management. Ongoing collaboration is essential in the rapid development of this field and this Workshop series continues to be a forum for proactive strategic planning.

REFERENCES:

1. Altorki N, Lane ME, Bauer T et al. Phase II proof-of-concept study of pazopanib monotherapy in treatment-naive patients with stage I/II resectable non-small-cell lung cancer.J Clin Oncol. 2010 28(19):3131-7.
2. Buckler, AJ, Mulshine JL, Gottlieb R, Zhao B, Mozley PD, Schwartz LS. The Use of Volumetric CT as an Imaging Biomarker in Lung Cancer, Acad Radiol. Jan;17(1):100-6, 2010.

3. van Klaveren RJ, Oudkerk M, Prokop M, Scholten ET, Nackaerts K, Vernhout R, et al. Management of lung nodules detected by volume CT scanning. *N Engl J Med*. 2009;361(December 3 (23)):2221–2229[PubMed PMID: 19955524]
4. National Lung Screening Trial Research Team. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening.N Engl J Med. 2011 Jun 29. PMID: 21714641
5. International Early Lung Cancer Action Program Investigators, Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, Miettinen OS. Survival of patients with stage I lung cancer detected on CT screening.N Engl J Med. 2006 Oct 26;355(17):1763-71. Erratum in: N Engl J Med. 2008 Aug 21;359(8):877. N Engl J Med. 2008 Apr 24;358(17):1862. N Engl J Med. 2008 Apr 24;358(17):1875. PMID: 17065637
6. Institute of Medicine. A foundation for evidence-driven practice: a rapid learning system for cancer care: workshop summary, 2010; Washington, DC: National Academies Press.
7. Mulshine JL, Baer TM, Avila RS. I Virtual Focus Issue: Imaging in diagnosis and treatment of lung cancer. Opt Express. 18(14):15242-15282, 2010. doi: 10.1364/OE.18.015242.
8. Kostis WJ, Yankelevitz DF, Reeves AP, Fluture SC, Henschke CI. Small pulmonary nodules: reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up CT. Radiology. 2004 May;231(2):446-52.