FDA-Project Data Sphere® Symposium VIII:

Artificial Intelligence in Tumor Imaging: Needs, Opportunities & Challenges

Symposium Summary

Table of Contents

I.   Introduction

Over 125 experts were convened on November 11, 2019 at Stanford University for a cross-disciplinary symposium on AI in Tumor Imaging. This is the eighth in a biannual series of topical symposia co-sponsored by Project Data Sphere and the U.S. Food and Drug Administration (FDA) to address specific issues and opportunities in biomedicine. Stanford University School of Medicine hosted the symposium.

Participants described the potential for AI to augment imaging in cancer screening, characterization, staging, and response assessment. Quantity and quality of data to support model development was the subject of much discussion. Datasets available in a single institution may not support development of robust, generalizable image processing models because of variation between institutions in patient demographics, imaging parameters, and imaging hardware. Imaging data collected in multicenter oncology clinical trials are very useful for model development because the data are comprehensive and well-characterized, and usually formatted in a manner that supports interoperability. Centralized and federated frameworks for AI model development in imaging processing and some of their respective requirements were discussed. Progress in automated image processing will require respect for the existing regulatory framework and the evidence base that supports it.

Multiple speakers emphasized the urgency to take immediate, practical steps to move the field forward. Dr. Demetri proposed that the community's initial focus in data sharing and image processing model development be CT data from prospective clinical trials which included imaging of lungs, including a common disease such as lung cancer and one uncommon disease such as sarcoma. Dr. Khozin suggested automation of secondary review as a pragmatic goal that can be achieved in the near future. Because the FDA does not set standards, and instead adopts consensus in the medical community that has reached a certain threshold when developing guidance, the onus is on the community to drive consensus standards for AI models in image processing.

Under the leadership of the Images & Algorithm Task Force and in collaboration with dedicated industry partners, the Project Data Sphere platform has recently been enhanced to integrate clinical trial imaging datasets. Although this initiative is in the early stages, imaging data from ~2000 patients will be available on the platform by the start of 2020. The community is welcomed to start utilizing the data and, just as importantly, contribute additional clinical trial imaging datasets.

II.   Meeting Summary by Session

SESSION 1:   Welcome

Lloyd Minor, MD, Dean, *School of Medicine, Stanford University*

Dean Minor welcomed everyone on behalf of the Stanford University School of Medicine, and recognized the partnership of the FDA and Project Data Sphere in organizing the symposium. He introduced a theme that resonated throughout the day of how health care has lagged other fields in adopting technology. Dr. Minor described the unique opportunities presented by the convergence of technological innovations from multiple fields in biomedicine today, and challenged everyone to continue to lead the way in translating these innovations into improvements for patients.

SESSION 2:   Call to Order

Martin Murphy, DMedSc, PhD, FASCO; Chief Executive Officer, *Project Data Sphere*

Dr. Murphy set the stage for the symposium in the context of the ongoing work of the Project Data Sphere initiative (www.projectdatasphere.org). Project Data Sphere was started in 2013 as an independent initiative of the CEO Roundtable on Cancer Inc.'s Life Sciences Consortium and launched the open-access data sharing platform in 2014. Since then, there have been 26 publications in top-tier journals derived from the data. There has never been a data breach and capabilities are continually expanding. Project Data Sphere, in partnership with the FDA, has held 8 symposia focused on specific issues and opportunities in biomedicine, bringing together experts from FDA, industry, and academia.
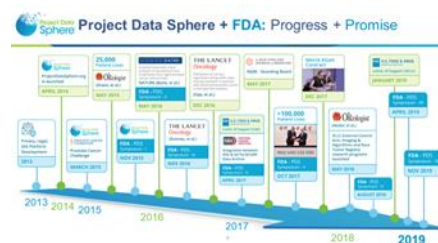
Dr. Murphy recognized Mace Rothenberg (Pfizer) and David Reese (Amgen), co-chairs of the Life Sciences Consortium, for their leadership. He extended sincere thanks to James Goodnight and colleagues at SAS for development and ongoing support of the Project Data Sphere platform and analytics. He shared his gratitude to past and current members of the FDA leadership, Robert Califf, Sean Khozin, Geoffrey Kim, and Richard Pazdur, for their partnership. He thanked Lloyd Minor for convening the symposium at Stanford and Reed Jobs (Emerson Collective) for his vision and encouragement. Dr. Murphy called on participants to be inspired by the day's program to pursue further collaborations and to always remember the patients—those who contributed to Project Data Sphere with their clinical trial data and those who will benefit from the biomedical discoveries made possible by the initiative.

SESSION 3: Life Sciences Consortium of the CEO Roundtable on Cancer

David Reese, MD; Executive Vice President, Research and Development, *Amgen*; Co-Chair, *Life Sciences Consortium*

Dr. Reese described the origin of the Life Sciences Consortium and the core of its mission, which is to be bold and venturesome in the fight against cancer. He shared a vision for the future in which data sharing and collaboration through Project Data Sphere expands to integrate additional types of data – imaging, molecular data, registry data, and real-world data. Ultimately the future lies in having all these data accessible and tractable from an analytics standpoint. A collaborative project with AACR to incorporate molecular data from Project GENIE into Project Data Sphere has begun. Dr. Reese described the imaging project as absolutely critical because it has the potential to be transformative for the conduct of clinical research.
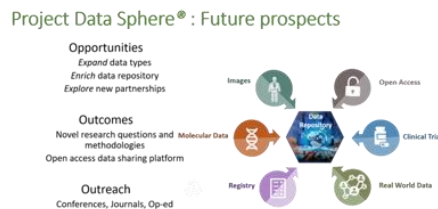
SESSION 4: Project Data Sphere: Convener, Collaborator, Catalyst

**Bill Louv, PhD; President, *Project Data Sphere***

Dr. Louv described the tremendous growth and productivity of the Project Data Sphere platform and programs. The platform serves as both a digital library and laboratory, with powerful SAS analytic tools available on the platform free for authorized users. New users generally obtain access to the whole population of data within 48 to 72 hours of applying. There is no requirement to submit a research protocol that must be approved by committee, as in a conventional gatekeeper access model. The database now represents >100,000 patient lives and 24 tumor types. Approximately 40% of the clinical trial data on the platform are from experimental arms. The platform is expanding, for example, with the integration of images and algorithms which is being led by George Demetri and Larry Schwartz, co-chairs of the Project Data Sphere Images & Algorithms Task Force. Recently, Project Data Sphere launched research programs to drive data acquisition in support of specific research needs in oncology. Dr. Louv emphasized that for the Project Data Sphere initiative to continue to catalyze progress in oncology, barriers to data sharing need to be addressed so that data acquisition can be accelerated.
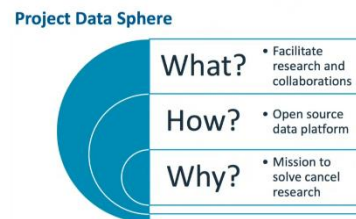
SESSION 5: Images & Algorithms Program Overview

George Demetri, MD; SVP, Experimental Therapeutics, *Dana-Farber Cancer Institute*; Professor and Co-Director, *Harvard Medical School and Ludwig Center at Harvard*; Co-Chair, *Images & Algorithms Task Force*

 Dr. Demetri described the pivotal role of imaging as a universal measure of response to anticancer treatments in oncology, both in clinical care and research trials. He pointed out that the interpretation of imaging differs in those settings, with use of RECIST largely confined to clinical trials. Additional rigor is achieved in clinical trials with blinded independent central review (BICR), however central radiology reads are typically slow,
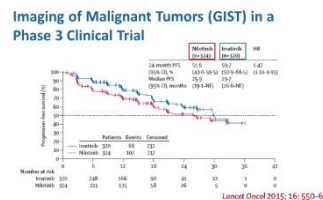
cumbersome, and expensive, yet still have risk of discordance. AI-based imaging interpretation has the potential to be more efficient, improve the consistency of interpretation, and even allow more information to be captured compared with assessment by any formal system such as RECIST. With deference to radiology colleagues, Dr. Demetri proposed that the community's initial focus, the "baby step," in data sharing and image processing model development be CT data from prospective clinical trials with imaging of the lungs, since the anatomy and contrasts in that location will likely be the most straightforward. He described other active issues that are being worked through for the Project Data Sphere Images & Algorithms program (patient privacy protection and secondary use consent, avoidance of misleading secondary analyses, how to structure access, how to link curated clinical data and external radiologic reads to imaging data, and cost) and encouraged everyone to think of what practical next steps are needed to drive the program forward.

SESSION 6:   Artificial Intelligence at the Cutting Edge of Imaging and Drug Development

Lawrence Schwartz, MD; Chairman, *Department of Radiology, Columbia University Medical Center***; Co-Chair,** *Images & Algorithms Task Force*

Dr. Schwartz expanded on the major role of imaging in oncology and the potentially transformative power of AI. Imaging has important roles in cancer detection, characterization, staging, and response assessment. Approximately half of all patient visits to oncology centers involve imaging, and 40%–45% of all imaging in radiology departments is cancer-related.
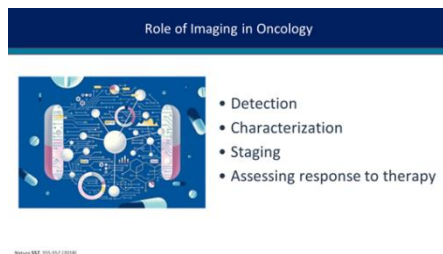
Dr. Schwartz shared vignettes that illustrate current gaps in tumor imaging that could be addressed with AI. Lung cancer screening with low dose CT is more effective then mammography, yet less than 5% of individuals at elevated risk for lung cancer are screened. The biggest barrier to screening is not patient education or reimbursement, rather it is the high false positive rate of current screening methods and resulting burden of office visits and repeat testing. AI models could help reduce this false positive rate. AI models may even help identify precancerous states that might be amenable to preventive therapy. With regard to tumor characterization, AI models could be used to generate tumor risk scores with enhanced predictive power. For example, Dr. Schwartz and colleagues are studying how AI models could integrate radiomics features and demographics features in patients with liver lesions to generate quantitative risk scores that provide additional information beyond Li-RADS classification. For tumor staging, PET-CT is generally a gold standard but access to PET scanners is limited globally. Enhanced information extraction from CT imaging would be useful. Dr. Schwartz and colleagues at Columbia, including Firas Ahmed, showed in a proof-of-concept study that a CNN could predict the SUV_{max} (a PET measurement that serves as a surrogate for malignancy) of lymph nodes based on CT images and primary tumor histology. In monitoring tumor changes over time (including response to treatment), AI-based imaging interpretation could permit more features to be quantitated than are assessed with current metrics (eg, RECIST) and may reveal clinically relevant radiomic biomarkers. Dr. Schwartz emphasized that for the potential of AI in imaging and drug development to be met, what is needed for the most part is annotated, curated data, and a way of sharing that data.

SESSION 7:   Fireside Chat: Evolution of Cancer Imaging Technology in Research

Robert Califf, MD; Advisor and Board Member, *Verily***; Vice Chancellor for Health Data Science,** *Duke University*
Lloyd Minor, MD, Dean, *School of Medicine, Stanford University*

Dr. Califf and Dr. Minor discussed how clinical research is evolving, and some opportunities and challenges associated with looking beyond large, randomized, controlled clinical trials to help answer clinical questions. They discussed areas in which the full potential of technology in health care has not yet been realized.

Dr. Minor: How do you see DCRI [Duke Clinical Research Institute] but also other comparable organizations moving from the world of randomized, controlled trials—still, by many viewed as the gold standard for some things? Moving now into the world where information can be derived from so many sources of data, oftentimes, in ways that you can't control and that you shouldn't try to control—how will that change the face of clinical research?

Dr. Califf: If you look at the new plan for the DCRI that's being rolled out now, it's totally built around being a hub for real-world evidence. And here you brought up the randomized trial. I think deeply embedded in the 21st Century Cures Bill and the publications from the FDA is a concept that we need to continuously learn using real-world data. But randomization and real-world evidence are not polar opposites. In fact, the best kind of real-world evidence is likely to be randomized real-world evidence.

And so, it's really critical to think of one dimension, which is where do you get your data. I think we're emerging from an era where the only way we can get credible data was to create a separate research universe, collect everything separately from clinical practice, which is very expensive and makes it very hard to do generalizable studies. Now we're moving into an era where data use of everyday is getting better and better. There are still many issues to work through.

Then the evidence is generated by applying a method to that data. I think we're going to ultimately realize that more often than people think, randomization is going to be the best method. Or, as I like to say, God's gift of randomization because it can help us avoid these errors that happened when we think we can control for all the confounding. I think we're all seeing a lot of big mistakes being made by people trying to apply causal inference to observational data—and getting the wrong answer. So we're all going to learn, there are going to be times to do both. We'll learn more and more from the observational, experiential data. But in many circumstances, we're going to need to say there is a crystallized question that we need the answer to and randomization up front is going to be the fastest way to get the answer.

Dr. Minor: It seems like we've cured cancer thousand-fold over in mice, as well as a variety of other diseases. There are some obvious reasons why we can't do everything in humans that we can do in animal studies. But what can we do to accelerate our knowledge and understanding of <u>human</u> biology?

Dr. Califf: It's been an issue where we were just limited by computation for what we could study in the human. Now with the kind of imaging that we're looking at today, the broad capability to store data and manipulate it, I think more and more of the human being will be the object of study. If you think about a human biologically, we're a system of systems. That is really hard as a problem. I think it explains a lot about why treatments fail, but they might work in a simple animal model, then you put it in the human

and they're all these redundant and off target systems that are all interacting at the same time, that are very unpredictable. But we can now begin to approach that computationally. To me, that's really the most exciting thing.

To me, the biggest question—I think, the purpose of this meeting, as I understand it, is how do we overcome our human cultural limitations to taking advantage of what's in front of us now?

SESSION 8: Spotlight Talks – Panel 1: Filling Current Gaps in Radiologic Image Processing

Moderator: Anshu Jain, MD; NCI/FDA Senior Fellow in Oncology Innovation, *National Cancer Institute*

Dr. Jain introduced the session with an overview of how imaging algorithms have evolved in sectors outside health care, including some of the advances driven by the ImageNet visual recognition challenge and a snapshot of data sharing within a federated learning model. Dr. Jain highlighted that to build AI models for tumor imaging in a way that will be reasonable and useful requires patient data that accurately represents the heterogeneity of the patient care experience in cancer.
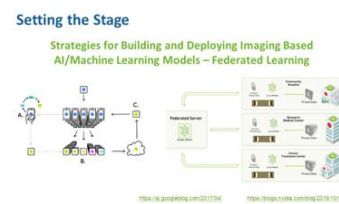
Curtis Langlotz, MD, PhD; Professor of Radiology and Biomedical Informatics, *Stanford University*

Dr. Langlotz outlined priorities for AI in medical imaging that were established by the 2018 NIH/RSNA/ACR/The Academy Workshop in which he participated.
Priorities for foundational research:
- Enhance raw images
- Automate image labeling and annotation
- Develop clinical decision support
- Develop methods for explaining information generated by AI algorithms to clinicians

Priorities for translational research:
- Encourage data sharing to ensure robust datasets for algorithm training, testing, and validation
- Establish standards for clinical integration of AI algorithms
- Create software use cases with common data elements
- Ensure a balanced regulatory framework

Dr. Langlotz described practical implementation of these priorities into the infrastructure and activities of Stanford's Center for AI in Medicine & Imaging. With regard to sharing data, Dr. Langlotz and colleagues see this as an obligation. Stanford has released 5 public datasets. One of these datasets included over 600,000 labeled chest radiographs, released jointly with MIT. Dr. Langlotz highlighted the need to improve standardization in imaging, to facilitate development of new quantitative tools for imaging interpretation. To this end, the Quantitative Imaging Biomarkers Alliance (QBIA) has developed imaging protocols that include detailed specifications, eg, for device calibration, patient prep, acquisition parameters, reconstruction, and resolution.
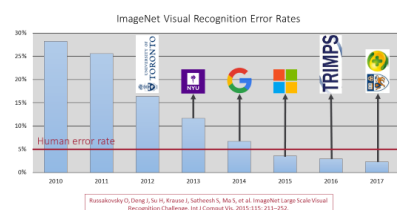
Daniel Rubin, MD; Professor of Biomedical Data Science, Radiology, and Medicine, *Stanford University*
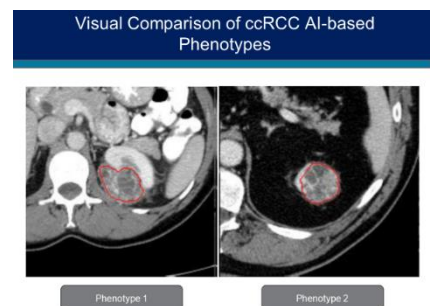

Dr. Rubin outlined opportunities for AI in cancer imaging including lesion detection, lesion segmentation, diagnosis, treatment selection, and response assessment. Lesion detection and segmentation are the most immediately amenable to AI-based automation. Automating RECIST would reduce variability and streamline assessments. Dr. Rubin shared an example from his lab of an AI model for segmentation of brain MRI developed as part of a larger research effort to fully automate brain lesion detection and segmentation. However, Dr. Rubin explained, datasets available in a single institution may not support development of robust, generalizable AI models for quantifying tumor burden because of variation between institutions in patient demographics, imaging parameters, and imaging hardware. Lack of generalizability will limit utility. Collecting datasets from multiple institutions/organizations in a central depository is one solution to the data bottleneck, however there may be challenges (eg, patient privacy issues, intellectual property considerations).
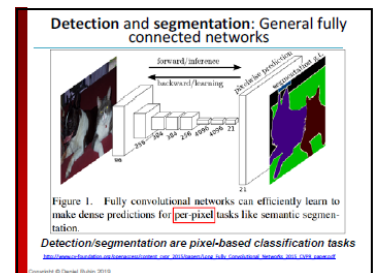
Federated learning models allow organizations to maintain data control; the algorithm learns locally and retains only revised parameters (not data). There are many models for federated learning and which model is optimal may depend on the particular task and distribution of data among sites (eg, Dr. Rubin and colleagues found that performance of a cyclical weight transfer model most closely matched performance of a model developed with centrally hosted data for a lesion classification task with data uniformly distributed among 4 sites). In addition to real-world data heterogeneity across sites, differences in computing hardware and network bandwidth among sites present challenges that need to be managed to make federated learning work. Dr. Rubin and colleagues are exploring optimizations to mitigate the adverse impact of data heterogeneity on federated learning model performance. In addition, there is evidence that the more sites included (more data) in developing a federated learning model, the closer accuracy approaches that of a model based on centrally hosted data. Dr. Rubin summarized that quality of care would go up dramatically if there were widely accessible algorithms to routinely measure lesions (eg, automation of RECIST). Collaboration and data sharing are key to accomplishing this.

Firas Ahmed, MD, MPH; Assistant Professor, Department of Radiology, *Columbia University Medical Center*

Dr. Ahmed shared a radiomics machine learning use case based on clear cell renal cell carcinoma CT data from The Cancer Genome Atlas and The Cancer Imaging Archive. The objective of the study was to identify a radiomics biomarker. The AI model discriminated 2 different phenotypes that correlated with different staging, grade, percent necrosis, post-resection recurrence risk, and cancer-specific survival. However, further investigation revealed that a technical parameter (slice thickness) was actually a predictor of the radiomics phenotype. Dr. Ahmed and colleagues went on to fine-tune the model to exclude radiomics and AI features that are affected by technical parameters and identified a non-enhancing component of the tumor as a prognostic biomarker for recurrence and cancer-specific mortality. Dr. Ahmed described how algorithms could be developed to serve as quality control systems to mitigate the confounding effect of technical parameters, using as an example a quality control

algorithm developed by Dr. Schwartz and colleagues to grade portal venous phase timing in liver CT. There are many technical parameters in tumor imaging that need to be considered and accounted for in AI model development, for example, by statistical regression, transfer learning, or simply starting with a homogeneous dataset.

Q&A

The panelists and audience members discussed the need for adequate data to support development of robust AI models for tumor imaging, and what kind of data are needed. They touched on considerations for integrating AI into radiology clinical workflow.

Dr. Jain: This discussion for federated learning, do you think that in terms of developing this type of collaboration [between multiple partners including pharma], it would require, could it be done in a purely federated manner? Or do you think it would be potentially a hybrid between training and building the model in a centralized fashion but then deploying it in a federated or distributed manner?

Dr. Langlotz: With respect to federated learning, I think it's a very powerful tool. I think that the concerns that Daniel raised are likely to continue to be a problem. If you look at the various sources of heterogeneity, they're going to lead to a need for data science expertise at each site to assure that the labeling process, for example, is similar, to assure that the image acquisition process is similar. Some of those we can overcome with just lots of data. But I think that a well-designed experiment or AI algorithm development process would have some thought to those areas of heterogeneity and try to control them in advance of the federated learning because I think that's going to make that federation much more powerful.

Dr. Rubin: I think there's going to be a synergy, you know, federated paradigm, as Curt said, is going to have its challenges to develop models that are as robust as centralized data, but I think that those challenges over time through research are potentially surmountable and highly synergistic with data that's available publicly. The important message is, we need to move forward on multiple fronts to make advances. There's been a lot of talk of us doing these things and not as much action. What holds us back is people engaging and stepping up to either make data available or to start entertaining participating in networks of federation. My plea to people in this room is try and get engaged. I think there's great opportunities if you do.

Dr. Califf (*Verily, Duke*): This is a great session. I wonder if you can comment on the issue of curated clinical outcome data to go with the imaging data because it seems like you're at risk of being like a boat over the horizon with no compass if you're just looking at images and you don't know what happened to the patients. Then related to that, it seems like one major advantage of federation is the data doesn't have to leave the institution. The rules that are in place for preventing re-identification by external parties could be enforced. I wonder if you could sort of talk about clinical data to match the imaging data and then the issues that you see in play related to privacy and re-identification.

Dr. Langlotz: Well, the issue you raise about institutions and privacy, I think, actually could be problematic even with federated learning. These neural network models could, in theory, memorize the data and could come away with some representations that would lead to the ability to re-identify that data. There are these homomorphic encryption and other methods being worked on. It's a complicated area and that's where I think some caution with respect to federated learning is warranted.

With respect to the outcome point you make, it's a very good one. I like to think of two classes of problems. If you have, let's say, a nodule in the lung, whether that nodule is present is going to be in the radiology report and that's going to be a reasonably useful way to label that study. In particular, you may have a case like congestive heart failure, where we're never really going to have a good standard short of a Swan-Ganz catheter, as to whether that patient truly has congestive heart failure. Going back to the lung nodule now, is that nodule cancer? Well, that's where we really do need to go back to the patient's electronic record. That's where we collaborate with people like Nigam Shah and others here at Stanford who are doing digital phenotyping and analyzing the narrative and structured information in the EMR to determine does this patient have cancer, when were they diagnosed with cancer. We can really use that as a reference point for what the images showed. Then on the flip side, you can use the images to try to predict what the outcome might be of that cancer diagnosis.

Dr. Demetri: I'd love everybody's opinion about whether there's data out there that already could be accessed for use or if we're talking about starting anew with some new standards so that we go forward in, let's say, a year. In a year, you get enough data to start this federation. Or do we go backwards? Do you think it already exists?

Dr. Langlotz: Federation always ends up being a topic of conversation because it's just so attractive. I think it really is a potential solution when you have multiple competing companies, each of which have data sources and they might work together in that way where they might not otherwise. I'll just come out and say it for the clinical world, if you look at what ImageNet has done for computer vision outside of medicine, we do not have anything equivalent in healthcare. We need a large pooled dataset of medical images that will spur progress in computer vision within medicine just as it has outside of medicine. I don't know how we get there, but I really think we need to get there. That has to do with how we manage the privacy issues and some of the IRB issues and other things at various institutions, but I really think that's an important goal.

Dr. Rubin: One distinction, just to keep in mind, is there's a huge difference between clinical trial data and real-world evidence. The real-world evidence—your question about EMR data—highly messy, inaccurate, transcription errors, poorly curated. But clinical trial data, high quality, especially FDA trials, audited, verified, validated, double checks, sometimes central reads, high quality. Imagine having that available for research and high volume in the kinds of problems we could really make inroads in. I mean, I'm a huge advocate of tapping into all of the EMR data but that takes us down the road of those messy issues you talked about. I think there's a lot of research to be done to be overcoming that, although, with weak learning and other related methods, computer science, there's potential value there. But I think if we could just get access to high quality of clinical trial data, add images to the data already available in this project, that would be huge.

Dr. Ahmed: I just wanted to reiterate what Dan was saying is basically we don't need to reinvent the wheel. The wheel has already been invented. Clinical trials have been done by pharmaceutical companies, curated metadata is available, and the identification of images are already been done. All that we need is to have a faith and cooperation, and come to a platform and collaborate with each other to take away the competition and just focus on, "Let's use this gold mine to build an AI." Instead of digging into the electronic medical record in the hospital and find some data that might introduce more noise than actual solution.

Dr. Langlotz: I wanted to make a comment about labeling. Everyone has their story similar to the confounders that Firas mentioned in his talk, where you find that slice thickness or something that really shouldn't matter ends up mattering. The laws of putting together a dataset that doesn't have those problems haven't been repealed, and they still apply even though it's, you know, computer scientists doing machine learning versus epidemiologist analyzing clinical trial data. But one thing that I think we do find over and over again is that poor quality labels can be overcome by large amounts of data for training. So these neural network models are relatively robust to error-prone labels. Look at that across, not just in healthcare, you know, lots of publications from Facebook and other places that show that you can introduce a large amount of error into labels and still these algorithms can reach outstanding performance. You still need, in healthcare, you need a very beautifully, highly accurately labeled test set to produce your ROC curve or whatever result plot you want to generate but that's a much smaller set of cases, typically in the hundreds. These low cost but error-prone data labeling methods are extremely powerful. We use those all the time to generate hundreds of thousands of cases.

SESSION 9: Spotlight Talks – Panel 2: Advancements in Image Recognition in Medical Imaging, Cancer Imaging, and Beyond

Moderator: Pratik Shah, PhD; Principal Investigator, *MIT Media Lab*

Dr. Shah began the session with a brief overview of how his translational computer science group is deploying machine learning systems in different areas of medicine. These include automated staining and destaining of tissue biopsies to achieve time/cost efficiencies and permit reuse of patient samples; use of a reinforcement learning system that learns from clinical trial dosing and de-escalation decisions to make individualized dosing recommendations; and analysis of real-world evidence to assess the efficacy of primary care and the ability of electronic health records to support informed decisions.
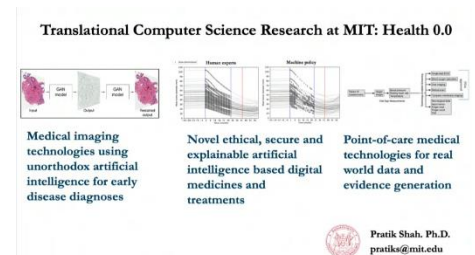
Courtney Ambrozic, MS; Senior Associate Staff Scientist, *SAS Institute, Inc.*

Ms. Ambrozic walked everyone through how she and her colleagues developed a deep learning model to fully automate liver lesion segmentation in 3D on CT imaging. The dataset encompassed 89 images from 42 patents with metastatic colorectal cancer. With the model, average DICE score=93% for liver segmentation and 66% for lesion segmentation (state-of-the-art DICE scores are 96% and 67%, respectively).

Shravya Shetty, MS; Senior Staff Engineer, *Google Health*

Ms. Shetty described how she and her colleagues developed a deep learning model to predict lung cancer risk based on low-dose CT imaging of the chest in patients at elevated risk for disease. Lung cancer screening by conventional methods is associated with a high rate of false positives (5%–13%) and false negatives (15%–21%). Dr. Shetty and colleagues used a subset of the National Lung Cancer Screening Trial (encompassing 15,000 patients, 44,300 cases, and 121,000 volumes) for training, tuning, and final validation. In addition, an independent dataset from an academic medical center was used exclusively for final validation. A model was developed that combined lesion detection and classification. The model achieved radiologist-level performance when CT images from more than one time point were available

for a given patient, and actually outperformed radiologists when CT images from only a single time point were available.

Gregory Goldmacher, MD, PhD, MBA; Executive Director, Translational Biomarkers, *Merck Research Laboratories*

Dr. Goldmacher discussed types of imaging data collected in clinical trials, how clinical trial data are stored, some opportunities for AI, and approaches to overcoming challenges to data sharing. Imaging data collected in oncology clinical trials are comprehensive and well-characterized, and usually formatted in a manner that supports interoperability. Typically all clinical trial imaging (for every timepoint) that has passed the quality control process is stored. The various tumor imaging-based response criteria (RECIST, RANO, Cheson) used in clinical trials share similar methodology in that all define a set of target lesions that are measured and describe other lesions qualitatively. Hence, stored images include outlines for target lesions and markups for non-target lesions. Also stored are information from the reads, including tumor locations and categories, measurements and qualitative judgments, and calculated responses for every scan. Imaging data are usually formatted and stored in conformity with CDISC standards. Older data might be in non-compatible formats. Imaging data are typically stored at the independent clinical research organization, but are controlled by the pharma sponsor.
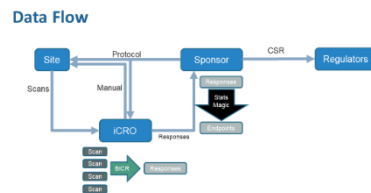
Possible sponsor concerns with data sharing include re-analysis of treatment efficacy, new safety questions, and imposed restriction of treatment eligibility (eg, requirement for companion diagnostic). How could data sharing be de-risked and made more efficient? Possible solutions are design of smart datasets that don't allow the full clinical trial dataset to be backed out, using trusted third parties to facilitate data sharing, employing standard contracts/agreements for data sharing, and maybe even help from regulators. Dr. Goldmacher shared some examples of how clinical trial imaging data have been used in-house at Merck to develop AI-based tumor imaging models (for tumor segmentation and to identify inflammation).

Tito Fojo, MD, PhD; Professor of Medicine, *Columbia University Medical Center*

Dr. Fojo discussed approaches to streamlining drug development through modeling. Dr. Fojo and colleagues used modeling to define benchmarks for exponential tumor growth in patients undergoing different standard-of-care treatments for colorectal cancer using radiographic data from Project Data Sphere. By modeling the pooled colorectal cancer data, they showed that tumor growth rate tertile (and octile) correlates with overall survival. Tumor growth rate of an experimental treatment could be benchmarked against standard treatments and might serve as a surrogate for clinical outcomes. In a comparison of volumetric, bidimensional, and unidimensional growth data (modeled based on data from a FOLFIRI + aflibercept study in colorectal cancer), volumetric data had the greatest power for discriminating between treatment groups and thus allowed for the smallest sample size. Dr. Fojo and colleagues undertook similar analyses in prostate cancer, using data from Project Data Sphere and PSA to model tumor growth rate with similar findings. Volumetric tumor growth rate data might have utility when making go/no go decisions based on
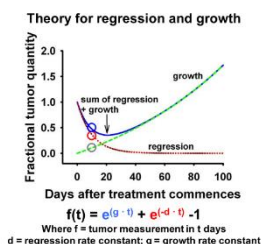
small datasets. Tumor growth rate modeling could help inform trial design to allow minimization of internal control arms or even allow for entirely external control arms.

Matt Lungren, MD, MPH; Associate Director, *Stanford Center for Artificial Intelligence in Medicine and Imaging, Stanford University*

Dr. Lungren discussed the precedent for open-source data catalyzing major advances in AI, eg, Deep Blue beating Garry Kasparov, Watson becoming the world Jeopardy Champion, and computer vision surpassing human recognition performance in the ImageNet challenge in 2015. In these examples, the key algorithm proposals were old; the open-source data used to train the algorithms were new. Dr. Lundgren advocated that once clinical data have been used to provide care, secondary use of that data should be for the benefit of all through open source research and education. To this end, Stanford researchers have established the Medical ImageNet, a multi-institution open repository of 5 million labeled images with plans to obtain half a billion images in the next five years. Over 7,200 teams worldwide competed in a bone x-ray imaging interpretation challenge hosted by Medical ImageNet. Alternative approaches to solve data availability challenges include possibly augmenting existing datasets with synthetic data in the future (eg, in rare diseases) and using federated learning systems. However, these alternative approaches will not have the same catalytic effect as having large amounts of open-source data available to the global research community.

Q&A

The panelists and audience members discussed many aspects of data sharing, including the role of the patient as an advocate for data sharing, overcoming barriers to sharing inside industry, managing data heterogeneity, and what constitutes responsible data sharing. The discussion touched on the future. including an expectation that the role of the radiologist will evolve (and certainly not become extinct) because of AI and a prediction that some forms of data sharing are likely to become mandatory, therefore proactive design of workable solutions is warranted.

Dr. Shah asked how patients could be engaged to advocate for and drive data sharing. Could patients simply share their de-identified data directly to Medical ImageNet and upload it to the cloud? Could patients ask the trial sponsor to release their data to Project Data Sphere?

Dr. Fojo shared the opinion that there isn't a patient who enrolled in a clinical trial who wouldn't agree to data sharing as long as their PHI was shielded. He highlighted the fact that Project Data Sphere has been operating for five and a half years, and nothing negative has come out of it. He suggested that clinical trial data sharing could be made automatic, for example, required 5 years after FDA approval of the drug that was supported by the trial data.

Dr. Goldmacher said there is precedent for patient advocacy groups driving consensus and alignment among different stakeholders. For example, a few years ago the National Brain Tumor Society got together the FDA, academics, and industry and obtained alignment on endpoints in brain tumor trials. This led to a consensus imaging protocol that was rolled out in journals and across the industry. It is now standard of care.

Dr. Reese (*Amgen*): Just a couple observations. Number one, I think most of the fears about data sharing are imagined and not real, and they haven't come to pass, and we have analogs. 20 or 30 years ago, you didn't have to submit the structure of a molecule when you submitted a manuscript. Now, you do. Everyone thought the sky would fall in when that requirement was put in place. Now, no one thinks about it. Now, you have to submit [*not audible*] data when you're publishing a genetics paper. Again, the sky hasn't fallen in, and in general, the fears have not been realized. I don't think there's any reason to believe that this will be any different.

Number two, I think if we project ourselves 10 or 15 years into the future, this is going to be mandatory. I mean, society is moving in a direction, governments are moving in a direction where this sort of data transparency and availability will be mandatory, so we can either be proactive or reactive. To me, that's the only question that faces us right now. We can design these systems upfront and offer something, or we can react to something that's imposed, so I would say let's approach it from that framework, and let's get it right from the start.

Dr. Eben Rosenthal (*Stanford*) posed a question about experience to date with pooled clinical trial data—is quality of imaging data generally the same across trials?

Dr. Lungren answered that heterogeneity across trials is something that can be addressed. Different institutions have their own scanner and imaging protocols. The model will account for resulting heterogeneity. The lack of standardized reporting in radiology is another source of heterogeneity, however if there are enough data, labelling noise doesn't matter. As long as the test set is clean, it doesn't matter. This has been demonstrated multiple times and is an argument for having a large dataset.

Dr. Shah agreed and highlighted that there is also a problem of misrepresentation and bias in race, gender, ethnicity in datasets. If there is an imbalance in the training dataset (eg, certain races, diseases, or phenotypes are underrepresented), this can result in bias in the algorithm.

Ms. Shetty also agreed and explained that she and her colleagues typically introduce noise in the images that they use, so that models are more robust to noise.

SESSION 10:   Keynote: Insights to Accelerate Progress in Assessment of Drug Effect – Including Cancer Imaging Technology

Robert Califf, MD; Advisor and Board Member, *Verily*; Vice Chancellor for Health Data Science, *Duke University*

Dr. Califf discussed the need for better approaches to health care in the US. Over the last 5 years, predicted life expectancy at birth in the US has declined. There are large disparities between counties in overall life expectancy and cause-specific mortality rates. Although potentially excess deaths from cancer have declined dramatically in urban counties, only modest declines are observed in rural counties. With these data in mind, Dr. Califf advocated that any new technologies in cancer, including imaging technologies, ideally be scalable and financially accessible to ensure they reach everyone who needs them. Technology will have the effect of democratizing health care and some changes are already underway. Consumers are accessing health care information directly; there are 1 billion health-related

searches on Google every day. Optimizing the quality of the health information retrieved is important. The extent of a patient's right to their data when they participate in a clinical trial is an open question. In medical imaging, AI models have great potential to improve the reliability of imaging measurements and increase the information content. Other areas of medicine, eg, pathology, will also be transformed by AI.

Progress can be accelerated when the evidence base for the existing regulatory framework is respected and considered in the development of new technologies. Dr. Califf cautioned that a distinction should be made between early translational research to identify a potential biomarker and the many steps involved in analytical validation, clinical validation, and fit-for-purpose validation. The last step, fit-for-purpose validation, entails qualification for use from a regulatory standpoint and demonstration of clinical utility (such as favorable cost benefit). Dr. Califf walked through the different kinds of biomarkers and encouraged everyone to consult the FDA's Biomarkers, EndpointS, and other Tools (BEST) resource for comprehensive information. With respect to biomarker categories, a common error is to consider a biomarker predictive of treatment effect when it is actually only prognostic of disease outcome, independent of treatment. This error may be made when only data from patients receiving experimental treatment are analyzed; a control arm is critical. Due to the complexity of human biology, there are multiple reasons why candidate surrogate endpoints may fail validation (Fleming & DeMets. *Ann Intern Med*. 1996). From a US regulatory standpoint, surrogate endpoints are characterized by the level of clinical validation. Only a validated surrogate endpoint can be used for full approval of a medical product. Sometimes a reasonably likely surrogate endpoint can be used for accelerated approval (a form of approval that requires a postmarketing confirmatory study be conducted). Dr. Califf cautioned that a correlate does not a surrogate make. Although many radiologic features will correlate with outcomes, that does not automatically mean that they can be surrogate endpoints. It will take a lot of discipline and rigor to follow through on the various aspects of validation. Dr. Califf closed by sharing how impressed he is with the progress that has already been made with AI in tumor imaging, and calling for more data sharing to allow the field to move quickly so that it can bring benefits to patients.

SESSION 11:   Expert Panel: Artificial Intelligence in Image Processing –Needs, Requirements, and Challenges

Moderator: Lawrence Schwartz, MD; Chairman, *Department of Radiology, Columbia University Medical Center***; Co-Chair,** *Images & Algorithms Task Force*

Panelists:
Tito Fojo, MD, PhD; Professor of Medicine, *Columbia University Medical Center*

Mace Rothenberg, MD; Chief Medical Officer, *Pfizer, Inc.*

Pratik Shah, PhD; Principal Investigator, *MIT Media Lab*

Sam Gambhir, MD, PhD; Chairman of the Department of Radiology, *Stanford University*

Karla Childers, MS; Senior Director, Office of the Chief Medical Officer, *Johnson & Johnson*

Q&A
The panelists had a wide-ranging discussion about the needs, requirements, and challenges of AI in image processing.

Issues surrounding patient consent for data sharing

Dr. Schwartz put forth the example of pixel-only imaging data, de-identified and with no associated metadata—would that be human subject data?

Ms. Childers commented that regardless of level of de-identification, the commitment made to the patient and the parameters to which the patient actually consented should be honored. Just because something is legal, does not mean it is ethical. It is incumbent on researchers to be good stewards of the data entrusted to them. Ms. Childers and colleagues are working on defining what is appropriate around secondary use of data and how to have conversations with candidate clinical trial participants around consent that are not coercive, for example, in the setting of a trial that is a last resort for a patient.

Dr. Gambhir concurred, "Even if it's that single pixel, or voxel to be more technically accurate, from a given individual, I think that causality chain back from the fact that it originated from a given human is important to consider in any downstream use of that data. I don't think it matters whether you've segmented it or are using a portion of the data. It comes back to then another belief that not every patient's the same in terms of their views of what they want done with their data or are willing to have done with their data…You have a more complex problem in that not each person would agree with what can be done with that data that originated from them. So it makes our jobs that much harder. But I believe that's the way in which you should approach any dataset or data element that you obtained from a human."

Dr. Shah wondered if patients would be more comfortable with and feel more enriched by sharing their data if they received feedback regarding how their data has helped advance research – for example, on a very simple dashboard.

Dr. Rothenberg posited that issues surrounding patient consent for data sharing might be more appropriately dealt with in a panel with greater patient advocate representation. "The idea is, and as Marty pointed out, we are dealing with data, which are the patients with the tears removed. This is something that they're passionate about. People have given their lives to generate these data. So, if we can communicate effectively to the patients, whether they be in clinical trials or in clinical practice and say in order to help, maybe not you, but your children and your grandchildren and many other people, who will be treated better and diagnosed earlier and have better outcomes because we're able to aggregate your data with thousands and millions of other patients—are you willing to do that? I'd be hard-pressed to find a single patient who would say no. So, I think there must be some way that we can engage people who really are going to be able to move the needle—patients and patient advocacy organizations to not only request but demand that this be done."

Possible barriers to data sharing in academic medical centers

The panel discussed why obtaining data from an academic medical center might be challenging. Although centers vary, barriers can include cost and limited resources (eg, for the work of de-identifying and curating), internal siloes and lack of a centralized process for releasing data, lack of an internal strategic plan for releasing data, and even ambiguity over which entity within a medical center owns the data.

Some experiences with data sharing by industry

Ms. Childers shared the experience at Johnson & Johnson with data sharing. Johnson & Johnson shares data from clinical trials of approved medical products within the Yale Open Data Access project (which has a gatekeeper model for access) and Project Data Sphere. Over 250 trials have been shared. There has only been one request for re-analysis, which ultimately confirmed the original trial findings. The advantage of the collaboration model, in which data sharing is done with a trusted third party on a secure platform, is that there is a standard process for data sharing that does not get bogged down in legal agreements.

Dr. Rothenberg shared some perspectives on data sharing specifically related to the investigation of safety signals. The sponsor wants to ensure they have a complete, up-to-date, and accurate safety database. Collaboration between the external investigator who may have preliminary evidence of a possible safety signal and the sponsor is critical to determining whether there is a real safety concern. "Because nobody wants to put something out there that is erroneous, only to be refuted years later." Dr. Rothenberg described the example of the smoking cessation drug, Chantix. Based on case reports of homicidal and suicidal ideation or actions, a Boxed Warning was added to the product labeling and FDA required a postmarketing study of 8000+ patients to evaluate neuropsychiatric safety. Usage dropped dramatically. Based on the safety profile of Chantix in the postmarketing study, FDA subsequently removed the Boxed Warning and usage increased. Thus an erroneous signal resulted in fewer patients benefitting from this method of smoking cessation.

Dr. Schwartz described some challenges he and colleagues experienced when trying to obtain clinical trial imaging data from certain industry sponsors. When asked to share data, sponsors typically said yes or maybe. A prompt no would be very useful because in at least once instance, the answer was ultimately no but in the interim Dr. Schwartz and colleagues spent years of time and effort to come to a data sharing agreement with that sponsor. Requesting specific datasets worked best. They focused on trials with centrally collected data. Enlisting help from the principal investigators of the original trials was useful for demonstrating that the new analyses were scientifically worthwhile and not going to be fishing expeditions. Asking for only the control arm data is one way to get initial agreement. A key learning was to not settle for a different, less useful dataset than the one requested and expending time and resources analyzing it.

Clinical trial data vs real-world data

Dr. Schwartz asked everyone to consider whether there is any framework/construct for comparing real-world data to clinical trial data so that as we start to look for imaging data to populate repositories, there is an understanding of the relative benefits versus costs of obtaining more real-world data. For example, is there a way to equate the value of a given clinical trial dataset with a given number of real-world clinical patients in terms of algorithm development?

How AI is changing functional imaging

Many of the day's presentations focused on CT and MRI. Dr. Schwartz invited **Dr. Gambhir**, as an international expert, to provide a perspective on functional imaging and novel tracer development.

Dr. Gambhir shared some history of functional imaging and described how regulatory simplification (approval pathway for tracers, which are administered in nonpharmacological doses, was separated from that of drugs) accelerated tracer development. Functional imaging is still underutilized in drug development and could be used more often to provide information about drug occupancy in a patient and predicted response to therapy. With AI, it is now possible to predict molecular imaging with one tracer based on molecular imaging obtained with a different tracer (eg, predict brain FDOPA-PET based on FDG-PET). Multiplex analyses incorporating molecular imaging plus other types of information (eg, blood biomarkers, imaging obtained with other modalities) to predict response or for cancer screening are now possible with the advent of AI.

Applying quantitative AI-based approaches in areas of medicine beyond radiology

The panel and audience members discussed how many other areas of medicine rely in part on human visual interpretation for decision-making such as pathology, ophthalmology, gastrointestinal endoscopy, and surgery. Arguably AI-based approaches could create efficiencies and increase information capture in all these areas. A cultural shift in medicine will be needed for the full potential of AI to be realized. In some aspects, medicine is artisanal. The gap between what can be done quantitatively and the acceptance of those quantitative techniques in clinical practice will need to be closed.

SESSION 12:   Call for Action & Next Steps

Sean Khozin, MD, MPH; Associate Director, *Oncology Center of Excellence, FDA*; **Member,** *Images & Algorithms Task Force*
Robert Califf, MD; Advisor and Board Member, *Verily*; **Vice Chancellor for Health Data Science,** *Duke University*

Dr. Khozin and Dr. Califf discussed next steps for incorporating AI into imaging. Dr. Khozin predicted that in the future the field will move beyond use of RECIST 1.1 as a gold standard. He recommended that a stepwise approach is the most pragmatic way to proceed with automation in tumor imaging. As a first step, in registrational studies the secondary radiology review in cases of discordance between initial readers could be automated. Currently, the overall rate of discordance is around 30%, and is higher in certain hard-to-measure tumor types like pancreatic or ovarian. Automation of the second review would achieve time/cost efficiencies and ultimately bring therapies to patients faster. To accomplish this, data are needed to train algorithms.

Dr. Califf observed that automated, digital assessments have the potential to encompass many more data points than traditional efficacy endpoints. This introduces complexity into the efficacy assessment in a clinical trial and may be a concern for some pharma sponsors. Dr. Califf wondered if there is something a regulator could do to address this potential obstacle?

Dr. Khozin advised that there are definitely certain things that can be done to create the incentives needed to move to a more holistic data-driven approach to investigating efficacy. He also pointed out that collecting more data could be beneficial for identifying subpopulations of patients for whom a drug

provides benefit in situations where there is not a significant efficacy benefit in the experimental arm overall.

Dr. Goldmacher (*Merck*) commented that full automation of central review by RECIST is a terrific idea, but could be nontrivial to accomplish because there is judgment involved in selecting lesions to follow. Downstream functions are more amenable to automation.

Dr. Reese (*Amgen*): I'd like to issue a call to action. I think Project Data Sphere is a great place to start. There are submissions that are just launching. J&J actually submitted the first dataset to the imaging project. Actually taught us a lot about the challenges of simply setting up the pipes and how you download these datasets, which is not trivial. Amgen has two datasets that are in flight—literally within days or a few weeks they'll be downloaded. And that'll be a couple thousand patients in total across all of those data sets. You can actually get started on some work. But we need lots more data than that. And so, I'm certainly willing to pledge my time as part of the Life Sciences Consortium and Amgen's commitment to this. I think coming out of here, we will have a call to action to get participation broadly across industry. But not only industry, we need NCI support, and we need academic support as well. In Project Data Sphere, the simple reality is it's been hardest to get clinical trials data from academic medical centers. That's been the single largest challenge. On the NCI once we got things ironed out in terms of process has been a phenomenal contributor.

Ms. Childers (*Johnson & Johnson*) commented that J&J learned a lot through the process of sharing imaging data with Project Data Sphere and are now building into their imaging management process considerations related to downstream sharing.

Dr. Khozin, in response to a question from the audience about the FDA's position on using synthetic clinical trial data to train algorithms for the initial steps of automating RESIST, clarified that the FDA generally cannot set standards. Instead, the FDA adopts guidance that has been developed by medical community consensus that reaches a threshold that it becomes acceptable to the FDA.

Dr. Rothenberg (*Pfizer*): Let me raise some ideas that we haven't really addressed so far today as other potential catalysts for changing this and moving field forward. So when we think about incorporating artificial intelligence into imaging and impacting the care and outcome for cancer patients, whether it be on clinical trials or in clinical care, how about if we introduce this approach into our radiology training programs? How about if we engage patients and make this an expectation rather than exception of the data they generate, whether it be in a clinical trial, or in clinical care? What about increasing funding for artificial intelligence and not only in radiology, but in multidimensional analysis of disease? Could we set standards and expectations, five-year action plans, by many of the stakeholders—academic medical centers, professional societies, health agencies, sponsors—for what we envision as the end state for this or the hoped-for state in five years, and how we're going to achieve that, to set as specific goals as we can. Lastly, how can we articulate the cost, both financial and to health outcomes, of not doing this, of our inaction. Those are some ideas. I'd love to hear people's reaction, including yours.

Dr. Califf: I think it's a pretty good list, and some of it does feed into what Sean said. In looking at what happened after 1962, what constitutes an adequate and well-controlled trial, what the statute says is "in the opinion of experts in the field." There's nothing stopping a group like this or friends of a group like

this, if enough momentum occurs, to create I won't call it a legal standard, but a consensus of experts in the field, which would then be adopted by FDA.

Dr. Khozin: I think this applies to a lot of other let's call them computational innovations that we're trying to bring to the table, that it's critical to, A, address the actual, practical problems, and not be too aspirational. When we talk about automating RECIST and the secondary review as the first step, that's not an aspirational goal. It's a very pragmatic goal, something that we can do today. Generating consensus around these themes and building that community consensus is a critical part of what we need to do over the next year or two.

Dr. Califf invited **Dr. Murphy** to deliver a charge to the group. Dr. Murphy thanked everyone for their participation in the symposium and dedication to the Project Data Sphere initiative. He issued a rousing call to action, "For all of you, but most of all for all the patients that you represent, one of whom was on the stage today, for them, we have a responsibility. No more dialogue. Get in there and let's act. If it's only just datasets, think how wonderful that's going to be."

SESSION 13:   Closing Remarks

Mace Rothenberg, MD; Chief Medical Officer, *Pfizer, Inc.*
Lloyd Minor, MD, Dean, *School of Medicine, Stanford University*
Reed Jobs, MA; Managing Director, Health, *Emerson Collective*

Dr. Rothenberg, **Dean Minor**, and **Mr. Jobs** extended thanks to everyone for their contributions to the symposium and commitment to seeing the area of AI in tumor imaging move forward and have a positive impact on patients.

Greg Simon, JD; Former President, *Biden Cancer Initiative*

Mr. Simon advocated that the organizing principle of health care is wrong because it does not center on what the patient wants and needs. Patients need to be empowered to manage their own health information and weigh in on their own health care. They need to routinely receive their medical information (analagous to statements issued by financial institutions to customers) and that information must be timely (eg, test results should be shared promptly). Patients want to support research. They need to be adequately informed about their rights with regard to sharing their data or tissue. It is possible to collect large amounts of data on a patient (eg, EHR data, genomics data, metabolic data, microbiome, and wearables data), but its value is diminished without the insight that can be obtained by speaking with the patient. The patient can provide context around data and give it greater meaning.

Mr. Simon proposed a patient-centered priority for AI in tumor imaging. He challenged researchers to work across disciplines to find new ways to detect the absence of tumor regression under treatment earlier with blood tests, even before it can be confirmed by imaging. This would allow patients to stop ineffective treatments earlier and limit their exposure to treatment-related adverse events. He advocated for evolving tumor classification with help from AI to encompass more radiomics features (imaging patterns) and be less anchored to location of a tumor in the body.

Mr. Simon: Do patients want to help you? Absolutely they want to help you. Are we letting them? No, and that's what we've got to fix. That's what we've got to fix, and that's what we can fix. This is the most uplifting and encouraging meeting I've been to on the subject of sharing data in my lifetime that wasn't

designed by the patients, but by the doctors. Thank you for showing up, and thank you for standing up, and let's keep it going. Thank you all.

Supplementary Information

BEST Resource (Biomarkers, EndpointS, and other Tools): Product of the Biomarker Working Group charged by the FDA-NIH Joint Leadership Council to develop a glossary of harmonized terminology for biomarkers and endpoints
http://www.ncbi.nlm.nih.gov/books/NBK326791/

Project Data Sphere
https://www.projectdatasphere.org

A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop
https://pubs.rsna.org/doi/pdf/10.1148/radiol.2019190613

A Road Map for Translational Research on Artificial Intelligence in Medical Imaging: From the 2018 National Institutes of Health/RSNA/ACR/The Academy Workshop
https://www.jacr.org/article/S1546-1440(19)30458-2/pdf

Abbreviations

AACR: American Association for Cancer Research

ACR: American College of Radiology

AI: artificial intelligence

BEST: Biomarkers, EndpointS, and other Tools

BICR: blinded independent central review

CDISC: Clinical Data Interchange Standards Consortium

CNN: convolutional neural network

CT: computed tomography

DCRI: Duke Clinical Research Institute

DICOM: Digital Imaging and Communications in Medicine

DICE: Dice similarity coefficient

EHR: electronic health record

FDA: Food and Drug Administration

GENIE: Genomics Evidence Neoplasia Information Exchange

IRB: institutional review board

NIH: National Institutes of Health

PHI: protected health information

RANO: response assessment in neuro-oncology criteria

RECIST: Response Evaluation Criteria in Solid Tumors

ROC: receiver operating characteristic

RSNA: Radiological Society of North America

SUV: maximum standardized uptake value

The Academy: The Academy for Radiology and Biomedical Imaging Research

FDA-PDS Symposium VIII

Artificial Intelligence in Tumor Imaging:
Needs, Opportunities & Challenges

Monday, November 11, 2019
7:30 AM – 5:15 PM

Li Ka Shing Center's Berg Hall
Stanford University School of Medicine
291 Campus Drive, Stanford, CA 94305

Hosted by: Stanford MEDICINE

EVENT AGENDA

Time	Session	Speaker(s)
7:30 – 8:00 AM	**Breakfast Buffet**	
8:00 – 8:10 AM	**Welcome & Call to Order**	Lloyd Minor, MD; Dean, School of Medicine, Stanford University Martin Murphy, DMedSc, PhD, FASCO; Chief Executive Officer, Project Data Sphere
8:10 – 8:20 AM	**Life Sciences Consortium of the CEO Roundtable on Cancer**	David Reese, MD; Executive Vice President, Research and Development, Amgen; Co-Chair, Life Sciences Consortium
8:20 – 8:30 AM	**Project Data Sphere: Convener, Collaborator, Catalyst**	Bill Louv, PhD; President, Project Data Sphere
8:30 – 9:00 AM	**Images & Algorithms Program Overview**	George Demetri, MD; Co-Director, Ludwig Center; Co-Chair, Images & Algorithms Task Force
9:00 – 9:30 AM	**Artificial Intelligence at the Cutting Edge of Imaging and Drug Development**	Lawrence Schwartz, MD; Chairman, Department of Radiology, Columbia University Medical Center; Co-Chair, Images & Algorithms Task Force
9:30 – 10:00 AM	**Break & Networking**	
10:00 – 10:30 AM	**Fireside Chat: Evolution of Cancer Imaging Technology in Research**	Robert Califf, MD; Advisor and Board Member, Verily; Vice Chancellor for Health Data Science, Duke University Lloyd Minor, MD; Dean, School of Medicine, Stanford University
10:30 – 11:45 AM	**Spotlight Talks - Panel 1: Filling Current Gaps in Radiologic Image Processing** *Current challenges and workflow needs for developing artificial intelligence techniques in image processing in a) image acquisition, b) image quality, c) image storage, d) regulatory and industry perspective*	Moderator: Anshu Jain, MD; NCI/FDA Senior Fellow in Oncology Innovation, National Cancer Institute Curtis Langlotz, MD, PhD; Professor of Radiology and Biomedical Informatics, Stanford University Daniel Rubin, MD; Professor of Biomedical Data Science, Radiology, and Medicine, Stanford University Firas Ahmed, MD, MPH; Assistant Professor, Department of Radiology, Columbia University Medical Center Q&A
11:45 AM – 12:45 PM	**Lunch Buffet & Networking**	

Time	Session	Speakers
12:45 – 2:30 PM	**Spotlight Talks - Panel 2: Advancements in Image Recognition in Medical Imaging, Cancer Imaging, and Beyond**	Moderator: Pratik Shah, PhD; Principal Investigator, MIT Media Lab Courtney Ambrozic, MS; Senior Associate Staff Scientist, SAS Institute, Inc. Shravya Shetty, MS; Senior Staff Engineer, Google Health Gregory Goldmacher, MD, PhD, MBA; Executive Director, Translational Biomarkers, Merck Research Laboratories Tito Fojo, MD, PhD; Professor of Medicine, Columbia University Medical Center Matt Lungren, MD, MPH; Associate Director, Stanford Center for Artificial Intelligence in Medicine and Imaging, Stanford University Q&A
2:30 – 2:45 PM	**Break & Networking**	
2:45 – 3:45 PM	**Keynote: Insights to Accelerate Progress in Assessment of Drug Effect – Including Cancer Imaging Technology**	Robert Califf, MD; Advisor and Board Member, Verily; Vice Chancellor for Health Data Science, Duke University
3:45 – 4:30 PM	**Expert Panel: Artificial Intelligence in Image Processing – Needs, Requirements, and Challenges**	Moderator: Lawrence Schwartz, MD; Chairman, Department of Radiology, Columbia University Medical Center; Co-Chair, Images & Algorithms Task Force Tito Fojo, MD, PhD; Professor of Medicine, Columbia University Medical Center Mace Rothenberg, MD; Chief Medical Officer, Pfizer, Inc. Pratik Shah, PhD; Principal Investigator, MIT Media Lab Sam Gambhir, MD, PhD; Chairman of the Department of Radiology, Stanford University Karla Childers, MS; Senior Director, Office of the Chief Medical Officer, Johnson & Johnson Q&A
4:30 – 5:00 PM	**Call for Action & Next Steps**	Sean Khozin, MD, MPH; Founding Director, Information Exchange and Data Transformation (INFORMED), Associate Director, Oncology Center of Excellence, FDA; Member, Images and Algorithms Task Force Robert Califf, MD; Advisor and Board Member, Verily; Vice Chancellor for Health Data Science, Duke University
5:00 – 5:15 PM	**Closing Remarks**	Mace Rothenberg, MD; Chief Medical Officer, Pfizer, Inc., Co-Chair, Life Sciences Consortium Lloyd Minor, MD; Dean, School of Medicine, Stanford University Reed Jobs, MA; Managing Director, Health, Emerson Collective Greg Simon, JD; Former President, Biden Cancer Initiative
5:30 – 7:30 PM	**Project Data Sphere Reception**	

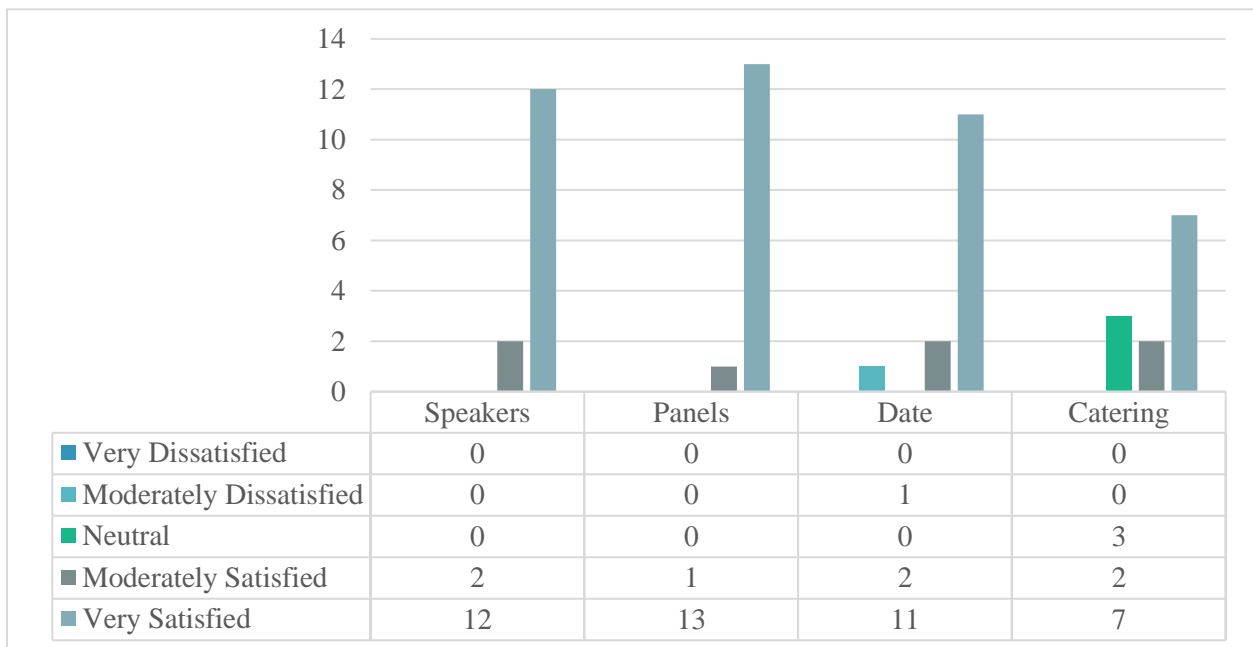Project Data Sphere® is an independent, not-for-profit initiative of CEO Roundtable on Cancer's Life Sciences Consortium
#FDAPDSsymp @ProjDataSphere | www.ProjectDataSphere.org

Twitter Metrics
@ProjDataSphere

28-day Summary	November 11–12 Summary
• 14,600 unique tweet impressions (↑**283%** compared to the previous 28-day period) • 449 profile visits (↑**532%**) • 94 mentions (↑**4600%**) • 45 followers gained	• 28 tweets resulted in **8200 organic impressions** • 25 retweets • 62 likes All of the above represent 1- and 2-day bests

Post-Event Survey Results

Satisfaction with Speakers, Panels, Date, and Catering

|	Speakers	Panels	Date	Catering
■ Very Dissatisfied	0	0	0	0
■ Moderately Dissatisfied	0	0	1	0
■ Neutral	0	0	0	3
■ Moderately Satisfied	2	1	2	2
■ Very Satisfied	12	13	11	7

14 out of 14 respondents plan to attend the Images & Algorithms Symposium in the future

- Reasons for intention to attend future symposia:
  - Important topic with much more to cover
  - To continue to learn from top experts in the field
  - Discussion needs to continue around merging clinical needs with tech advances
- Suggestions for ways to improve future symposia:
  - Industry reps from device (manufacturing) companies
  - More time for panel questions from audience
  - Progress report around Symposium VIII action items
  - Picture-in-picture webcast capability (ie, showing speaker and slide simultaneously)
 - Better breakfast food

Most liked about the symposium

- Mix of domain specialists and data scientists
- Fresh, modern feel of an academic research campus
- Well organized
- Advanced thinking
- Crowd size was conducive to networking
- And simply, but deservedly…Dr. Robert Califf

Likelihood of recommending symposium to a friend or colleague
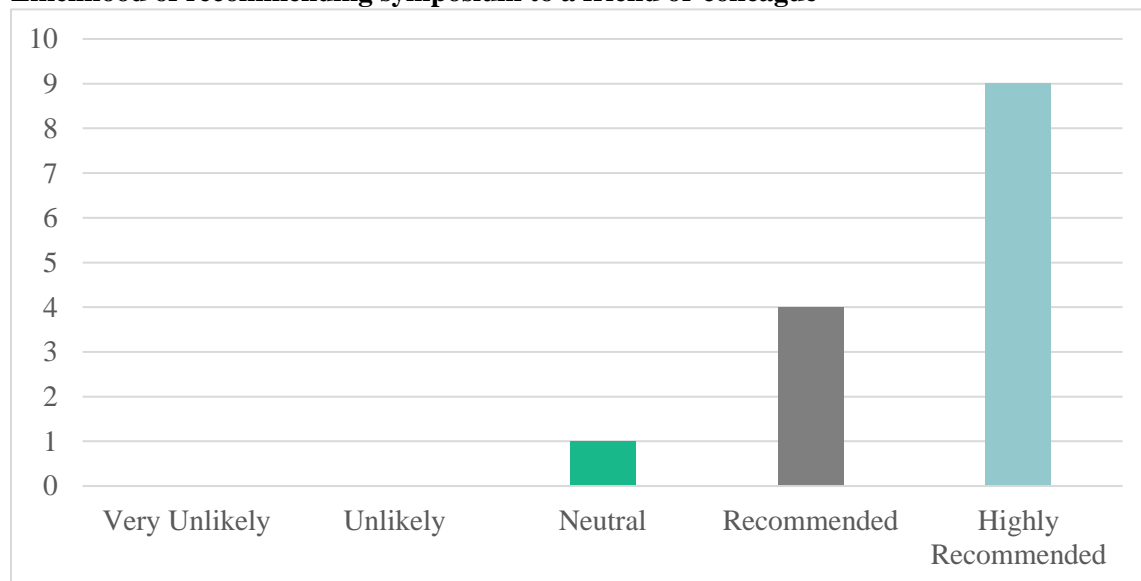
Satisfaction with the quality of networking opportunities
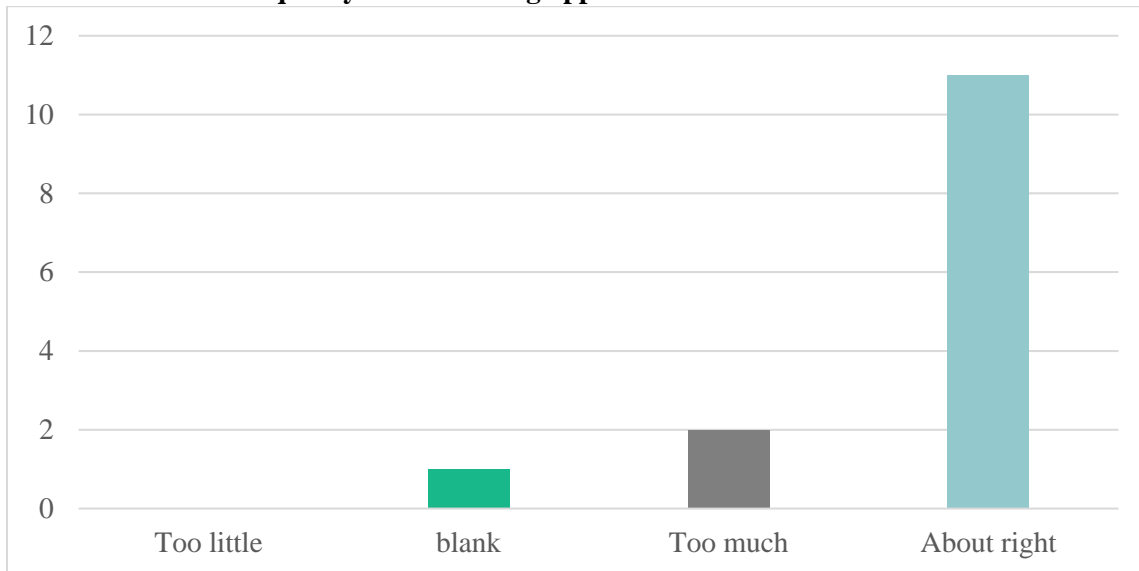
Affiliations of survey respondents:
- 5 academia
- 3 pharma/industry
- 4 non-profit
- 2 government

If you did not have a chance to provide feedback, you may submit your responses to the post-event survey at: **https://www.ceo-lsc.org/survey/fda-pds-symposium-viii**

Symposium Participants by Organization

Amgen	1	Massachusetts General Hospital	1
Biden Cancer Initiative	1	Merck Research Laboratories	1
CCS Associates, Inc.	2	MIT Media Lab	1
CEO Roundtable on Cancer	1	National Cancer Institute	1
Columbia University Medical Center	3	Novartis AG	1
Dana-Farber Cancer Institute and Harvard Medical School	1	Paige AI	1
Duke University/Verily	1	Palo Alto Medical Foundation Research Institute	1
Elektra Labs	1	Pfizer, Inc.	1
EMD Serono	1	Project Data Sphere	8
Emerson Collective	5	PVmed Technology	1
Evercore	1	SAS Institute, Inc.	3
Google Health	1	Stanford University	47
HealthTech	1	Syapse	3
Independent Consultants	2	UC Santa Cruz	1
Invicro	1	US Food and Drug Administration (additional FDA staff attended remotely)	1
Johnson & Johnson	1	xCures	2
		Total	**98**