Innovations in Lung Imaging

Expanding the Possibilities for Nodule Detection and Tracking with Artificial Intelligence and Cloud Computing
Summary

Lung cancer is the leading cause of cancer death in the United States and worldwide.\(^1\) According to the American Cancer Society (ACS), in 2019 an estimated 230,000 people in the United States will develop lung cancer, and 143,000 will die from the disease.\(^2\) Globally in 2018, the two main forms of the disease—small cell and non-small cell lung cancer—accounted for 2.1 million new cancer diagnoses and 1.8 million deaths.\(^3\) Cigarette smoking is the primary risk factor for lung cancer, and in the U.S. smoking can be attributed to up to 90 percent of lung cancer fatalities.\(^4\)

Given how lethal lung cancer can be, the U.S. Preventive Services Task Force (USPSTF) recommends that those at high risk for the disease undergo annual screening tests.\(^5\) Testing is conducted using low-dose computed tomography (LDCT), and early detection—while the disease is potentially curable—has been shown to reduce mortality significantly. Today, according to the American Lung Association (ALA), around 8 million Americans are at high risk for lung cancer and are therefore prime candidates for low-dose CT scans. The ALA estimates that if even half that population underwent screening as recommended, at least 12,000 lung cancer deaths could be prevented.\(^6\)

Primary care physicians and other providers, as well as organizations like the ACS and ALA, have taken it upon themselves to spread the word about LDCT and the impact it can have in the fight against lung cancer. Still, there are a number of significant limitations associated with this technology that must be addressed if it is to meet its potential. While actual scanning, for example, only takes minutes, interpreting the images is a manual and time-consuming process, and efficacy depends on the skills of individual radiologists. Pulmonary nodules, when they are found, must be evaluated based on size, density, and other factors, and they must be monitored over time through comparison to other images from follow-up scans. Nodules can be extremely difficult to distinguish, and errors and missed diagnoses are common.\(^7\)

What follows is an overview of lung imaging innovations that use artificial intelligence (AI) and cloud computing to facilitate nodule detection and tracking in conjunction with LDCT. For radiologists and their support teams, this report will provide a snapshot of the latest in imaging technology, including tools that promise to not only augment workflow, but also improve clinical efficacy; and for anyone at risk of developing lung cancer, it should offer some level of hope that lung imaging can enable outcomes previously thought impossible.

---

2. World Health Organization. Cancer. Available at: [https://www.who.int/news-room/fact-sheets/detail/cancer](https://www.who.int/news-room/fact-sheets/detail/cancer)
8. NEJM Journal Watch. "Lung Cancer Screening in Real World Has High False-Positive Rate." Available at: [https://www.jwatch.org/fw112499/2017/01/31/lung-cancer-screening-real-world-has-high-false-positive](https://www.jwatch.org/fw112499/2017/01/31/lung-cancer-screening-real-world-has-high-false-positive)
Thoracic CT: A Key to Lung Cancer Survival

Lung cancer screening using low-dose computed tomography is advised over imaging via conventional chest X-ray for its higher sensitivity and specificity, and because it results in less radiation exposure for patients. During an LDCT procedure (“low-dose” because it uses less ionizing radiation than a conventional CT scan), the patient is exposed to a precise beam of X-rays that is rapidly rotated around their body. The thousands of signals produced by this energy are recorded and processed by a computer into cross-sectional “slices.” These two-dimensional slices are then digitally stacked to create three-dimensional images of the body’s internal structures.9

LDCT has been the only recommended screening test for lung cancer since 2013, when the USPSTF published its “Final Recommendation Statement” in the Annals of Internal Medicine.10 When the screening process reveals one or more lung nodules or lesions, follow-up LDCT scans are typically conducted over a period of months to monitor for changes in size. In some cases, lung nodules may be discovered incidentally via conventional chest CT scan during non-screening procedures.

In 2017, the Fleischner Society, the international medical society for thoracic radiology, published guidelines for the management of incidentally discovered nodules. Among the society’s recommendations: When such nodules are larger than 6 mm in diameter, follow-up scans should be conducted in all patients; while nodules smaller than 6 mm warrant “optional” follow-up scans for patients deemed at higher risk.11 Multiple studies have shown that early detection by LDCT can reduce lung cancer mortality by 20 to 43 percent among high-risk populations.12

9. National Institute of Biomedical Imaging and Bioengineering. “What is a computed tomography (CT) scan?” Available at: https://www.nibib.nih.gov/science-education/science-topics/computed-tomography-


LDCT Imaging Challenges

Once an LDCT scan is complete, the images are typically reviewed and interpreted by a radiologist. Cross-sectional slices can be displayed individually or as composite three-dimensional images. 3D images can be rotated in space; and when suspicious nodules are identified, cross sections can be analyzed manually—slice by slice—to determine their exact locations.

The interpretation and reporting process for suspicious findings has been described by the American College of Radiology in its “Practice Parameter” on the topic:\(^\text{13}\)

> “Lung nodules and focal lung lesions should be reported with respect to anatomic location (lung lobe, segment) and series/image number to facilitate comparison to both prior and subsequent thoracic CT examinations. Nodules should be described with respect to size, attenuation (soft tissue, type of calcification, fat), opacity (solid, ground glass [also known as nonsolid], and part-solid, containing both solid and ground-glass components), and margins (eg., smooth, lobulated, spiculated).”

The ACR also recommends that radiologists compare their findings in current images with those from imaging studies conducted previously in order to note the presence of any visible changes:

> “When comparing changes in nodule size, opacity, and contour, efforts should be made to compare the oldest scans available in addition to the most recent prior scan to assess for changes over time, including subtle changes.”

And finally, the ACR recommends that radiologists use its Lung CT Screening Reporting and Data System (Lung-RADS) to report LDCT exam results in a structured manner and “facilitate data collection and monitoring of patient outcomes.”\(^\text{14}\)

The process—even when assisted, as the ACR recommends, by standard computer workstation analysis—is tedious, time consuming, and skill-dependent. Inter-reader variability among clinicians means that sometimes nodules are marked and tracked accurately, and

---

13. The American College of Radiology. ACR-STR Practice Parameter for the Performance and Reporting of Lung Cancer Screening Thoracic Computed Tomography (CT). Available at: https://www.acr.org/-/media/ACR/Files/Practice-Parameters/CT-LungCsScr.pdf?la=en

sometimes they are not. Furthermore, as demand for scans has increased over the years, radiologists have been subjected to ever-increasing workloads, which in turn has led to concerns about their ability to keep up.

One 2015 study at the Mayo Clinic, published in the journal Academic Radiology, found that the typical radiologist in 1999 was required to interpret 2.8 CT scan images per minute. By 2010, the study revealed, that number had grown to more than 19 images per minute. “The average radiologist interpreting CT…examinations must now interpret one image every 3-4 seconds in an 8-hour workday to meet workload demands,” the researchers said. An article in Applied Radiology commenting on the study summed up the absurdity of the situation in its title, “The Radiologist's Gerbil Wheel.” The number of CT exams had gone up by 68 percent, and the average CT exam increased from 82 images in the first year of the study to 679 images in the last. “There has been little done to mitigate the impact of increases in imaging content on workload,” the Mayo Clinic authors concluded. “The effect of increased examination content on fatigue and interpretation accuracy remains a relatively undefined clinical problem and merits additional investigation.”

These findings warrant further concern given the fact that currently only a fraction of those at high risk for lung cancer actually undergo screening with LDCT (less than 5%, according to a 2019 study in the American Journal of Preventive Medicine). If more people receive screening, as the USPSTF recommends, the workload for radiologists may increase even more.

---


16. Applied Radiology. "The radiologist's gerbil wheel: interpreting images every 3-4 seconds eight hours a day at Mayo Clinic.” Available at: https://appliedradiology.com/articles/the-radiologist-s-gerbil-wheel-interpreting-images-every-3-4-seconds-eight-hours-a-day-at-mayo-clinic


Arterys Medical Imaging Cloud AI (MICA)

The Arterys Medical Imaging Cloud AI platform provides collective intelligence by:

• Aggregating data from around the world.
• Monitoring user input that can be used for further training.
Artificial Intelligence and Applications to CT

The good news for radiologists and the healthcare organizations where they work is that while the need for image analysis is increasing all the time, the capacity for computer programs to help is climbing as well. Computed tomography, as the name implies, has since its introduction in the early 1970s relied on computing algorithms to instantly produce three-dimensional anatomical images. Similarly, image analysis has for many years depended on the use of computer-based teleradiology and picture archiving and communication systems (PACS). Today, however, these digital tools are just the beginning, especially with the emergence of new technologies deploying artificial intelligence and machine learning.

“AI and its offshoots, machine learning and deep learning, are already changing radiology,” noted a 2017 article in Radiology Business. Radiologists will never be replaced by AI, the author predicted. But AI-based tools will enable them to work more efficiently and accurately, for example, by pre-analyzing images to “separate truly urgent items on image-interpretation worklists from those that can wait… while also performing routine reading tasks such as quantification, segmentation and pure pattern recognition.”

Other experts on the subject have reached similar conclusions. A 2018 article in European Radiology Experimental, for instance, spelled out the challenges faced by radiologists and described AI as a “tremendous opportunity” for improving the field: “With an irreversible increase in the amount of data and the possibility to use AI to identify findings either detectable or not by the human eye, radiology is now moving from a subjective perceptual skill to a more objective science.” AI algorithms mimic “natural intelligence” by looking at medical images “to identify patterns after being trained using vast numbers of examinations and images,” the authors explained. “Those systems will be able to give information about the characterisation of abnormal findings, mostly in terms of conditional probabilities to be applied to Bayesian decision-making.”

“Artificial intelligence,” an umbrella term coined by the computer scientist John McCarthy in the mid-1950s, applies to a field of computer science in which machines are developed to perform tasks

---


19. European Radiology Experimental. “Artificial Intelligence in Medical Imaging: Threat or Opportunity?” Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6199205/
that previously required human intelligence. Computer systems that utilize AI are often said to mimic human cognition in their capacity for learning, problem solving, and other activities. “Machine learning,” which is part of AI, was defined by another pioneer in the field—Arthur Samuel, in 1959—as “the ability to learn without being explicitly programmed.” “Deep learning,” in turn, is an approach to machine learning generally described as relying on computer architecture in which multiple processors are interconnected—much like the connections between neurons in the human brain. So-called “neural networks” can “learn” by a process of trial and error, and thereby become “smarter” over time to perform better and better on a given task.

A 2018 article in PLOS Medicine, “Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study,” explained the value of deep learning in the context of medical imaging: “In lieu of the often subjective visual assessment of images by trained clinicians, deep learning automatically identifies complex patterns in data and hence provides evaluations in a quantitative manner,” the authors wrote. Deep learning networks, they continued, “allow for the automated quantification and selection of the most robust features, and thus they require little to no human input.” That study, among other things, considered the ability for deep learning networks to characterize lung tumor characteristics seen on CT images compared to other technologies designed to do the same; and it evaluated their capacity for stratifying patients according to their mortality risk. Among patients who received surgery to remove cancerous nodules, deep learning technologies “significantly outperformed models based on predefined tumor features as well as tumor volume and maximum diameter,” the authors found. And deep learning networks also improved upon current prognostication methods for lung-surgery recipients, “hinting at their utility in patient stratification and potentially sparing low mortality risk groups from adjuvant chemotherapy.”

Improving Inter-Radiologist Agreement

To investigate whether reads augmented by AI detection improve inter-radiologist agreement, Arterys designed a multi-reader retrospective study of 30 thoracic CTs from two European hospitals.

Three selection categories were defined based on clinical reports: screening examinations with nodules, screening examinations without nodules and other pathology/incidentals with unknown nodules.

Within each category, five studies were randomly selected from each institution. Three radiologists (from three different institutions) read all studies twice—once with AI detection, and once without. For reads without detection, radiologists were asked to add nodules with a nodule auto-segmentation tool. For those reads conducted with AI detection, detected nodules were reviewed by the radiologist and either accepted or rejected. Missed nodules were added using the nodule auto-segmentation tool, and each radiologist tracked the number of added nodules and deleted nodules (for detection only) on a spreadsheet.

Results

Using AI reduced the variation in total number of reported locations of concern. It also increased total number of reported nodules and the number of studies reported to have at least one nodule.

Differences of 4x were observed among radiologists when not using AI compared to 2.8x when using AI. For the number of studies reported containing nodules, the observed differences among radiologists was 2.4x when not using AI compared to 1.8x when using AI. Percentage agreement between radiologist pairs (defined as both radiologists reported at least one nodule in the study, or both agreed no nodules were present) was a low as 67 percent when not using AI, and as high as 87 percent when using AI.

References

Take a look at the following white paper to read the details and see the graphics:

Bridging the gap: How Arterys Lung AI improves inter-radiologist agreement
Deep Learning and Nodule Detection

In the context of LDCT screening and other imaging applications, deep learning technologies offer radiologists unprecedented ability to automatically detect and segment lung nodules. The Arterys Lung AICT solution, for example, was trained, validated, and tested using an open dataset from the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) of detected and segmented nodules from more than 1,000 thoracic CT exams. With that data as its foundation, the system employs a diagnostic pipeline consisting of three connected, interacting models: a nodule proposal system, a nodule classification system, and a nodule segmentation system. Working together, these systems provide detection and segmentation of solid, semi-solid, and ground-glass nodules, with segmentation including long axis, short axis, volume, and average Hounsfield units. Nodules are listed in order of priority, and can be easily custom-sorted by the reader.

One recent study of the Arterys system found that across a population representative of screening and incidental findings, for nodules between 4 mm and 30 mm in size, it was able to detect 90 percent of true nodules (those found by at least 3 out of 4 radiologists), 83 percent of ground-glass nodules, and 100 percent of malignant-looking nodules. And for detected nodules and user-found nodules, the system provided volumetric segmentation as accurately as an expert annotator. Among the results: For nodules greater than 6 mm in diameter, the system's median relative volume error was 11.5 percent, compared to 21 percent for annotators; while for nodules smaller than 6 mm in diameter, the median relative volume error was 16.4 percent (versus 24.5 percent for the annotator cohort).

Similarly, the same study found that a deep-learning system can significantly reduce missed reader detections. Four blinded reads per scan were collected for 100 screening exams and 50 more exams for the diagnosis of pulmonary embolism. The results for each annotating radiologist (the "holdout annotator") were compared to the 2/3 consensus of the other three radiologists. The system was found to cut the rate of missed detections for the annotating radiologist by 60 to 70 percent for the screening population and by 42 to 66 percent for an incidental-findings population. Overall, the study found, an AI-enabled system increases the sensitivity of the radiologist, allowing them to monitor smaller and earlier-stage nodules that would otherwise be extremely difficult to detect.

In another retrospective analysis of 30 studies from two European institutions using stratified random sampling to represent both screening and incidental patient populations, researchers measured the impact of deep-learning-assisted detection on inter-radiologist

Metrics Matter

When it comes to lung cancer diagnosis and treatment, it’s imperative to have accurate quantification describing the progression of lung nodules over time. The challenge is that creating volumetric masks of nodules can be a tedious and time-consuming process, so there’s risk that accuracy may be compromised.

To evaluate the volumetric error of the 3D lung nodule-segmentation model available in Arterys Lung AI, a study was conducted comparing it to radiologist-created annotations. Researchers randomly selected 69 scans containing 126 nodules (marked by at least three out of four radiologists) from the Lung Image Database Consortium (LIDC). The ground truth volume of a nodule was determined by taking the average of the volumes derived from 3D manual segmentations of each annotator. With that ground truth volume, accuracy of both the individual radiologist and the AI model could be evaluated.

Results

For nodules greater than 6 mm, the median inter-rater relative absolute volume error (IRAVE) for radiologists was 21 percent, compared to a model relative absolute volume error (MRAVE) of 11.5 percent. For nodules smaller than 6 mm, median IRAVE was 24.5 percent, while the MRAVE value was 16.4 percent. In all cases, the AI model produced a smaller median error when compared to the variability between radiologists. Using Arterys Lung AI can provide radiologists with robust and reproducible nodule metrics for the diagnosis and tracking of lung cancer.

References

Take a look at the following white paper to read the details and see the graphics:

As Accurate as the Radiologist: How Arterys Lung AI provides robust nodule volumes
Reducing False Negatives

Missed lung nodule detections during CT reads are common among radiologists. To determine whether artificial intelligence can reduce incidence of missed detections, a study was conducted comparing reads augmented with Arterys Lung AI with those of individual radiologists who didn’t use the Arterys system.

Researchers collected four blinded reads per scan for 150 thoracic CT examinations. The studies were divided into 100 biopsy-confirmed lung cancer LDCT examinations (screening) and 50 pulmonary embolism examinations (incidental). Radiologists were asked to annotate all lung nodules in the studies, then annotations from individual radiologists were compared to those of the three other radiologists. Total number of false negatives for each radiologist was determined against a two-thirds consensus from the counterpart radiologist.

To measure the impact of the Arterys system, radiologist reads were augmented using the additional nodule detections. The number of missed detections was then reevaluated.

Results

Total number of annotated lung nodules for each radiologist ranged from 4.8 to 7.6 nodules per scan on the screening population, and between 1.8 and 2.5 nodules per scan for incidental findings. Total number of false positives per scan ranged from 1.1 to 1.8 for screening, and from 0.5 to 1.2 for incidental findings. Total false negatives ranged from 89 to 190 for screening and from 21 to 33 for incidental findings. For the screening population, the rate of missed detections ranged from 60 to 70 percent with 1.6 additional detections per scan; while for the incidental population, the reduced missed detections ranged from 42 to 66 percent with 1.6 additional detections per scan.

References

Take a look at the following white paper to read the details and see the graphics:

AI has your back: How Arterys Lung AI reduces false negative rates

variability. Three radiologists from separate hospitals read all studies with and without detection with an appropriate cooling-off period between reads to reduce study bias. When the researchers measured agreement for the presence or absence of nodules in a study, they found that: reader agreement improved for all cases when using the system for detection; pairwise reader agreement was as low as 67 percent without detection, but as high as 87 percent with detection; and that pairwise reader agreement improvement ranged from 3 to 10 percent when using detection.

Deep Learning and Workflow, Tracking, and Reporting

Deep learning technologies can also drastically improve radiology workflow and facilitate the tracking and reporting of nodules over time. In one retrospective analysis of 30 studies from four public and private data sources and institutions (chosen to represent screening and incidental patient populations), researchers measured workflow efficiency when using the Arterys system. Among the four radiologists from different hospitals who read all images with and without detection, the AI-enabled detection model reduced reading times by an average of 45 percent.

When it comes to longitudinal tracking, the Arterys system permits automated tracking across multiple exams, includes graphical displays of nodule progression, and offers a comparison table with key metrics of nodule change over time. A dedicated Lung-RADS scoring panel for screening examinations automatically captures key features and calculates scores, and reporting is automated as well—and can be customized to display relevant information.
Cloud Computing: The Key to Advanced Imaging Analytics

For AI-enabled imaging technologies to support clinical workflow, they must be fast, accurate, and accessible. But high-speed computing and data analytics require significant processing power, an expensive proposition for some healthcare organizations.

The solution for many in the industry has involved relatively affordable cloud services and infrastructure—the hardware and software that enable cloud computing. Thanks to its specialized architecture, cloud computing can provide organizations with unlimited processing resources on an as-needed basis. It’s exactly what’s required for advanced imaging analytics, which may depend on data sets of fluctuating size.

The scalability of cloud architecture and its distributed-computing and virtualization capabilities allow for consistent processing performance without the need for additional IT investments. Cloud computing also permits organizations with multiple sites to simplify and streamline collaboration, since data are available to any provider with a secure network connection. One recent survey found that 86 percent of healthcare providers in the United States are already using cloud-based services, while 75 percent of providers globally are planning to do the same.23 The worldwide market for cloud computing in healthcare is expected to climb from approximately $20 billion in 2017 to around $35 billion in 2022.24


Conclusion

Advanced lung imaging technologies, including cloud-based solutions employing artificial intelligence, have become essential to the ongoing fight against lung cancer. Lung cancer prevention will always be centered primarily around efforts to reduce smoking, but among the high-risk populations where the disease is most prevalent, diagnostic and clinical decision-support tools can improve the odds for long-term survival. AI-enabled technologies are allowing radiologists to detect and track suspicious lung nodules efficiently and effectively, and to collaborate with their colleagues in other disciplines. These tools may not eliminate lung cancer, but they can bring greater objectivity to its treatment and—in the process—improve patient care.