**Title:** The IASLC Early Lung Imaging Confederation (ELIC) Open-Source Deep Learning and Quantitative Measurement Initiative

Stephen Lam[a], MD, Murry W. Wynes[b], PhD, Casey Connolly[c], MPH, Kazuto Ashizawa[d], MD, PhD, Sukhinder Atkar-Khattra[e], BSc., Chandra P. Belani[f], MD, Domenic DiNatale[g], Claudia I. Henschke[h], MD, PhD, Bruno Hochhegger[i], MD, PhD, Claudio Jacomelli[j], Małgorzata Jelitto[k], MD, Artit Jirapatnakul[l], PhD, Karen L. Kelly[m], MD, Karthik Krishnan[n], MS, Takeshi Kobayashi[o], MD, PhD, Jacqueline Logan[p], BaN, Juliane Mattos[q], PhD, John Mayo[r], MD, Annette McWilliams[s], MBBS, FRACP, Tetsuya Mitsudomi[t], MD, Ugo Pastorino[u], MD, Joanna Polańska[v], PhD, Witold Rzyman[w], MD, PhD, Ricardo Sales dos Santos[x], MD, PhD, Giorgio V. Scagliotti[y], MD, PhD, Heather Wakelee[z], MD, David F. Yankelevitz[aa], MD, John K. Field[ab], PhD, FRCPath, James L. Mulshine[ac], MD, Ricardo Avila[ab], MS

[a] Department of Integrative Oncology, The British Columbia Cancer Research Institute and Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada
[b] International Association for the Study of Lung Cancer, Denver, Colorado, US
[c] International Association for the Study of Lung Cancer, Denver, Colorado, US
[d] Department of Clinical Oncology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan
[e] Department of Integrative Oncology, British Columbia Cancer Research Institute, Vancouver, British Columbia, Canada
[f] Department of Medicine, Penn State College of Medicine, Hershey, Pennsylvania, US
[g] Intellitech Innovations, Accumetra, LLC, Clifton Park, US
[h] Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, New York, US
[i] Clinical Professor of Radiology, Department of Radiology, University of Florida, Gainesville, US
[j] Data management, National Cancer Institute of Milan, Milan, Italy
[k] Radiology Department, Medical University of Gdańsk, Poland
[l] Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, New York, US
[m] International Association for the Study of Lung Cancer, Denver, Colorado, US
[n] Independent Consultant, Accumetra, LLC, Clifton Park, US
[o] Department of Diagnostic and Interventional Radiology, Ishikawa Prefectural Central Hospital, Kanazawa, Ishikawa, Japan
[p] Fiona Stanley Hospital, Perth, Western Australia, Australia
[q] Federal University of Health Sciences of Porto Alegre, Porto Alegre, Brazil
[r] Department of Radiology, Vancouver General Hospital and the University of British Columbia, Vancouver, British Columbia, Canada
[s] Fiona Stanley Hospital, University of Western Australia, Perth, Western Australia, Australia
[t] Department of Surgery, Division of Thoracic Surgery, Kindai University Faculty of Medicine, Osaka-Sayama, Japan
[u] Department of Surgery, Section of Thoracic Surgery, National Cancer Institute of Milan, Milan, Italy
[v] Department of Data Science and Engineering, Silesian University of Technology, Gliwice, Poland
[w] Department of Thoracic Surgery, Medical University of Gdańsk, Gdańsk, Poland
[x] Hospital Cárdio Pulmonar da Bahia, Hospital Israelita Albert Einstein, São Paulo, Brazil
[y] Department of Oncology, University of Torino, Torino, Italy
[z] Stanford Cancer Institute, Stanford University, Stanford, California, US
[aa] Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, New York, US
[ab] Roy Castle Lung Cancer Research Programme, The University of Liverpool, Department of Molecular and Clinical Cancer Medicine, Liverpool, UK
[ac] Internal Medicine, Graduate College, Rush University Medical Center, Chicago, Illinois
[ab] Accumetra, LLC, Clifton Park, US

**Keywords:** lung cancer screening, artificial intelligence, deep learning, emphysema, nodule detection, nodule volume measurement

**Funding:** The International Association for the Study of Lung Cancer and Astra Zeneca provided funding for this work. AstraZeneca played no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

**Disclosure Statement of Conflict of Interest:** This will be added once all COIs are collected.

**ABSTRACT (250 words, max 250)**

**Background**

With global adoption of CT lung cancer screening, there is increasing interest to use artificial intelligence (AI) deep learning methods to improve the clinical management process. To enable AI research using an open source, cloud-based, globally distributed, screening CT imaging dataset and computational environment that are compliant with the most stringent international privacy regulations that also protects the intellectual properties of researchers, the International Association of the Study of Lung Cancer (IASLC) sponsored development of the Early Lung Imaging Confederation (ELIC) resource in 2018. The objective of this report is to describe the updated capabilities of ELIC and illustrate how this resource can be utilized for clinically relevant AI research.

**Methods**

In this second Phase of the initiative, metadata and screening CT scans from two time points were collected from 100 screening participants in seven countries. An automated deep learning AI lung segmentation algorithm, automated quantitative emphysema metrics, and a quantitative lung nodule volume measurement algorithm were run on these scans.

**Results**

A total of 1,394 CTs were collected from 697 participants. The LAV950 quantitative emphysema metric was found to be potentially useful in distinguishing lung cancer from benign cases using a slice thickness ≥ 2.5 mm. Lung nodule volume change measurements had better sensitivity and specificity for classifying malignant from benign lung nodules when applied to solid lung nodules from high quality CT scans.

**Conclusion**

These initial experiments demonstrated that ELIC can support deep learning AI and quantitative imaging analyses on diverse and globally distributed cloud-based datasets.

**INTRODUCTION (3687 words)**

Lung cancer is the leading cause of cancer deaths globally with over 1.8 million annual deaths (1). At 22%, the five-year net survival for lung cancer is among the lowest of all types of cancer (2). Low-dose computed tomography (LDCT) has demonstrated 20-39% of lung cancer mortality reduction (3-9). This has led to the national roll out of LDCT screening globally (10). To optimize the delivery of lung cancer screening, there is increasing interest to use artificial intelligence (AI) deep learning methods to discriminate malignant versus benign lung nodules, determine biological behavior to personalize screening intervals and improve risk prediction among people with negative screens (i.e., with no or only small nodules) (11-16). These studies are usually limited to relatively small numbers of cases aggregated from screening participants at a regional level. To date, the limited availability of large collections of high-quality screening cases may not fully reflect the actual diversity of screening participants and types of CT scanners deployed globally for the lung cancer screening process.

To address the need of an open source, screening CT imaging dataset for AI research that meets the increasingly stringent privacy regulations and the need to protect the intellectual properties of researchers, the International Association for the Study of Lung Cancer (IASLC) initiated the Early Lung Imaging Confederation (ELIC) project in 2018 to develop a globally distributed and cloud-based database and computational resource for AI research (17). ELIC was designed to allow regional lung cancer screening programs to securely make their anonymized lung cancer screening CT scans and metadata available for computational analysis by global researchers and algorithm developers without transferring the data outside their region. Contributors to this resource are provided tools so they can make their screening participants' clinical information de-identified and their CT scans can only be quantitatively analyzed to generate analysis results within the secure ELIC environment; the CT scans themselves cannot be downloaded or used for other purposes. Similarly, ELIC is designed to provide a strict assurance to algorithm developers that their algorithms and algorithm results are secure and cannot be obtained or used by other parties. Other federated medical imaging data and computational environments exist (18,19), but they do not provide both assurances. Some of the existing resources such as the National Lung Screening Trial repository or The Lung

Image Database Consortium image collection (LIDC-IDRI) (https://ieee-dataport.org/documents/lung-image-database-consortium-image-collection-lidc-idri) are easily accessible and offer researchers the opportunity to download data collections and freely test their algorithms. However, some of the CT scanners and image acquisition parameters are not consistent with current standards. The location of the malignant nodules may not be clearly annotated, nodule follow-up data are not available and clinical data may not be continuously updated. There is a need for continuously updated data due to changing lung cancer screening entrance criteria, CT scanner technology, and population exposures over time. The intent of ELIC is to provide a continuously updated set of high-quality CT lung cancer screening datasets that will be useful to lung cancer researchers and AI developers for algorithm performance evaluation.

To illustrate how this globally distributed CT screening imaging resource can be utilized to advance AI research, we applied both deep learning AI algorithms and traditional image processing algorithms to the ELIC dataset.

**METHODS**

ELIC Structure

ELIC uses a hub and spoke architecture such that each clinical site uploads anonymized data to a spoke cloud instance. This cloud instance is a virtual computer located in their local region. Algorithm developers can upload their AI or quantitative analysis algorithms to the ELIC hub cloud instance which is distributed to spokes to run against spoke data when an experiment is started. When ELIC starts an experiment, the algorithm, encapsulated in a Docker virtual machine, is sent to each spoke and the spoke cloud instance then spawns local cloud instances (typically 5 to 10) to run the algorithm on the datasets. This approach allows ELIC to analyze spoke data collections with two levels of parallelization, one across the spoke instances and another across any number of local cloud computing instances. Another benefit of this approach is that ELIC can easily scale at any time to perform very large and demanding computational analyses that would be extremely difficult to achieve if actual computing hardware were assembled for an analysis.

A large number of advances have been made to ELIC since our initial description of the first phase of the project (17) resulting in a fully functional and useful cloud based ELIC environment. First, an open-source de-identification and data curation tool (20) was modified to support the specific data collection and quality control requirements of ELIC. This had the benefit of helping sites more quickly and easily prepare data in a common format and also catching data collection issues early in the curation process before the data was uploaded to ELIC. Numerous additional enhancements were made to make ELIC more maintainable, efficient, and secure. To better support deep learning AI developers, ELIC was modified to allow for using GPU processing, which is commonly used to accelerate deep learning algorithms.

In this second phase of the initiative, ELIC was populated with anonymized lung cancer screening cases from seven globally distributed institutions located in Vancouver, Ishikawa, Milan, Porto Alegre, Perth, Gdansk, and New York. Institutional Review Board approval was obtained at each spoke site for these anonymized cases. Each site was asked to provide 100 lung cancer screening cases at two time points where 25 cases were confirmed early lung cancers, 25 cases each had no nodules and no cancer, and 50 cases each with at least one non-malignant nodule. The second scan for lung cancer cases had to be obtained prior to any treatment. Each CT lung cancer screening case was provided with metadata describing the case including the image acquisition protocol, the coordinates of lung nodules, the diameter of each nodule, and the nodule type (solid, sub-solid or part-solid). The lung cancer status of the nodule at the time of data submission to ELIC as determined by the donating institutions was used as the source of clinical ground truth for all analyses reported in this manuscript. In addition, basic patient demographic information such as age, sex, smoking status, were also provided. All data was represented with nomenclature and data values consistent with the data dictionary used by the open source VAPALS-ELCAP lung cancer screening management system (21).

Image analysis

Automated analysis of the ELIC LDCT lung cancer screening cases was performed with the following deep learning AI methods and quantitative imaging algorithms:

A. The presence of emphysema on LDCT has been reported as an independent risk factor for lung cancer (22). To evaluate the potential for quantitative CT emphysema metrics to help differentiate lung cancer cases from benign cases, we evaluated the ratio of quantitative measurement values (mean HU, median HU, Std. Dev. HU, Perc15 (23), LAV950 (24)) between lung cancer cases and non-lung cancer cases within the lungs of the participants in the ELIC dataset. A higher ratio signifies that the lung cancer cases had a higher metric value than the non-lung cancer cases whereas a lower metric value was produced when the lung cancer cases had a lower metric value than the non-lung cancer cases. A measurement ratio of near 1.0 signifies that the emphysema metric did not differ between the two types of cases. The goal of this analysis is to identify which emphysema metric was better associated with an increased risk for the eventual development of lung cancer.

To achieve this, we first applied a deep learning AI CT lung segmentation algorithm that produced lung and lobe segmentation masks. **Figure 1** shows the result of overlaying lobe segmentation masks over the lung region in an ELIC case. This deep learning lung segmentation algorithm was accelerated by using the GPU processing capabilities available on Amazon Web Services cloud servers.

The AI lung segmentation algorithm in this study used PyTorch (25) and nnUNet (26), two widely used deep learning frameworks. The network uses deep supervision, instance normalization, and a LeakyReLU. Images are predicted with a sliding window using overlapping patches with a 50% overlap and gaussian weighted softmax aggregation in the overlap regions. The algorithm was run on spoke spawned AWS cloud server instances with GPU hardware capabilities.

B. A semi-automated quantitative CT lung nodule volume algorithm derived from the open source Accumetra Lesion Sizing Toolkit (ACM-LSTK) (27) was run on all solid lung nodule locations provided the radiologist at each site. This algorithm was designed to support volume measurement of small solid lung nodules with a minimum diameter of 5.0 mm that met the Quantitative Imaging Biomarkers Alliance (QIBA) CT small lung nodule profile requirement such as a slice thickness of ≤ 1.25mm (28). The volume of the primary lung nodule for each case was independently evaluated at each of two time points and subtracted yielding a volumetric

change in $mm^3$ and as a percentage of the volume of the nodule at the first time point. The QIBA minimum nodule growth table based on nodule size was used to determine if a lung nodule had grown beyond what can be accounted for due to imaging and software measurement error (27). As we had previously found that quantitative assessment of small nodules improved with the rigor of the image acquisition quality, we stratified the initial CT images collected in ELIC based on acquisition paraments to evaluate if this variability affected the outcomes of the AI evaluations.

Data Subsets

The clinical demographic information of the lung cancer and non-lung cancer cases are shown in **Table 1**. **Figure 2** provides a flow diagram showing how the lung cancer cases and non-cancer cases were used to evaluate a deep learning AI lung segmentation and emphysema measurement analysis and a quantitative lung nodule volume change algorithm. The AI lung lobe segmentation algorithm was run against 168 lung cancer cases and 529 non-cancer cases resulting in successful lung segmentation and measurement results for 139 lung cancer cases and 482 non-cancer cases (data subset 1). The cases excluded from data subset 1 had DICOM CT image data issues preventing 3D analysis with the algorithms we used such as DICOM CT image series that had changing CT slice thicknesses and missing CT slices. Most AI and quantitative imaging algorithms require a regularly spaced rectilinear array of CT HU data values on which to operate. Thus, having gaps and changes in the input data generally prevents automated analysis for most AI and quantitative imaging algorithms.

To create data subset 2 suitable for assessing the performance of a solid lung nodule volume algorithm, more data exclusions were needed:

- Cases where CT scan slice thickness > 1.25 mm
- Cases where CT scan slice spacing > CT scan slice thickness
- Cases consisted of part-solid and nonsolid lung nodules
- Cases where there was a problem with running on the data

After applying these exclusions, data set 2 included 44 (26.2%) lung cancer cases and 134 (25.3%) non-cancer cases.

To further evaluate the performance of semi-automated lung nodule volume change algorithm on the highest CT image quality data in ELIC, we further eliminated the following cases:

- Cases where CT slice thickness > 1.0 mm
- Cases where reconstruction kernel edge enhancement (28) was very high
- Cases where scanning interval was < 120 days
- Cases where very poor segmentation results were encountered.

This resulted in a data subset 3, consisting of 18 (12.9%) lung cancer cases and 63 (11.9%) non-cancer cases to assess the performance of the semi-automated solid lung nodule volume change algorithm with high CT image quality data.

**RESULTS**

In the 697 CT screening cases, 44.3% of these cases were female and 15.5% were from individuals who have never smoked, 40.5% had smoked in the past, and 44.1% were still smoking. The mean and standard deviation for age, pack years, and number of lung nodules was 62.8 ± 8.3, 38.1 ±27.8, and 1.3 ±1.3, respectively.

The AI lung lobe segmentation algorithm was run on all ELIC cases resulting in successful lung lobe segmentation and emphysema measurement results for 139 lung cancer cases and 482 non-cancer cases (data subset 1). Analysis of the emphysema metric data collected indicated that LAV950 applied to thicker slice thickness data had the greatest potential to separate lung cancer cases from non-lung cancer cases, among the emphysema metrics tested. **Table 2** lists the spokes from lowest slice thickness to highest slice thickness and the measurement ratios for all the emphysema metrics evaluated. This preliminary study data shows that LAV950 performed better than all other metrics in scans with a slice thicknesses > 1.25 mm.

**Table 3** shows the mean volume change and the Coefficient of Variation (COV) for benign versus malignant solid lung nodules in data subset 2. Benign nodules either showed minimal volume change or decrease in volume in the follow-up scan while malignant nodules

showed a large increase in volume. The volumetric measurements had a small coefficient of variation.

**Figure 3** shows an example of a semi-automated volumetric segmentation of a benign solid lung nodule scanned with 1.0 mm slice thickness and spacing at two time points. The radiologist measured diameter increased from 9.0 mm to 10.7 mm, which would be considered a large enough growth to indicate suspicion for lung cancer using diameter measurements and the lung-RADS requirement of growth >= 1.5 mm. However, the measured volume of the solid lung nodule changed by 13.5% (100 x (286.2 – 252.1)/286.2). When we evaluate a 13.5% observed increase in volumetric size of this 9.0 mm solid lung nodule against the 44.3% growth required by the QIBA small lung nodule profile volumetric change table, we found that the observed change in volume was not large enough to confirm growth beyond CT imaging and software measurement error. Thus, in this case, radiologist diameter measurement indicated the potential for this nodule being malignant and the volumetric algorithm correctly classified this nodule as not achieving a size change beyond imaging and software measurement error.

We then compared diameter and volumetric change measurements to classify lung nodule malignancy risk using the 44 lung cancer cases and 134 non-lung cancer cases (data subset 2). The sensitivity and specificity of classifying malignant nodules using semi-automated volumetric change analysis was 75% and 92%, respectively compared to 75% sensitivity and 98% specificity using the lung-RADS criteria of a ≥ 1.5mm diameter change with radiologists measured diameter measurements.

When the performance of volumetric nodule changes to classify lung cancer cases from benign cases using the image collection with the highest acquisition quality and excluded the incorrect volumetric segmentations (data subset 3), this resulted in a volumetric change lung cancer classification sensitivity and specificity performance of 89% and 100%, respectively compared to 75% and 98% with diameter measurement (**Figure 4)**

**DISCUSSION**

Here we presented the first AI and quantitative imaging analysis pilot study results from the new open-source, globally distributed, and cloud-based ELIC lung cancer screening imaging database and computational environment. ELIC operates entirely on a globally distributed cloud with high levels of security provisions ensuring that both clinical sites and algorithm developers can operate with high levels of control over their data and algorithms.

For this study ELIC was populated with 697 LDCT lung cancer screening cases from 7 global locations. Both a deep learning AI lung segmentation algorithm and a solid lung nodule volume measurement algorithm were evaluated against the ELIC data using the AWS cloud services infrastructure. This is proof-of-concept study demonstrates the capability to run complex deep learning and quantitative imaging analyses on diverse and globally distributed datasets without the need for transmitting the clinical images and metadata to a central location and with high levels of security compliant with existing international regulatory requirements. This cloud environment operates with a high levels of data security federating large globally diverse clinical/CT image databases. This computational resource enables the participation of many clinical sites and algorithm AI developers from across the globe, who would otherwise be unable to participate in such public health research or quality-controlled environment.

The idea of using federated data analysis technology to provide a secure environment to access large volume of high-quality imaging and biomedical data have been developed and used to link medical centers within a country to facilitate AI research (18, 19, 29, 30). However, to our knowledge these federated image collections were not from multiple countries and did not have as high a level of security as ELIC. For example, most federated systems utilize computing hardware running within a local healthcare setting, opening up the possibility that the computing hardware, along with the stored data and algorithms, can be stolen or accessed without permission by local administrators or others. This type of approach will require more computational staff support from many globally distributed computer system administrators and the attendant cost may be prohibitive in some settings. In contrast, ELIC data is stored on a highly secure cloud infrastructure (e.g. AWS) and managed by a trusted third party administrator, greatly limiting the total number of system administrators that have access to

the ELIC data. From a donor or national perspective, ELIC can also enforce strict data usage restrictions on the amount and type of data that can be extracted from an analysis by groups that run algorithms against ELIC data. Such capabilities include limiting the size of data analysis reports, the size of uploaded images, and applying lossy jpeg image compression on algorithm generated images to prevent unapproved copying of the image collections.

For this study we performed a preliminary evaluation of both an AI deep learning driven method for emphysema measurement and a quantitative imaging algorithm for measuring change in small solid lung nodules. Both analyses yielded promising preliminary results that have the potential to help inform the design of larger studies. The quantitative analysis of emphysema using a deep learning AI lung segmentation found that the LAV950 metric performed best at separating lung cancer cases from-non-lung cancer cases, but only when the site used thick CT slices. This finding is consistent with the lung density methods and results reported by Gallardo-Estrella (31) where thin CT image slices were combined to create CT scans with 3mm slice thickness. Given the lack of consensus as the best lung densitometry method is best to use (e.g., Perc15 or LAV950) to quantitatively assessing emphysema (32), the preliminary results found in this study merit further evaluation. It should be noted that the preliminary emphysema results reported here do not imply that CT scanning for quantitative assessment of emphysema should be performed with thick (>= 2.0 mm) CT image slices. Rather, these results may support that CT scans should be acquired with thin CT slice thickness and quantitative analysis algorithms should combine the thin CT section information to create voxels over a larger 3D region to help suppress image noise as previously reported (31).

This pilot evaluation also provided interesting findings regarding the performance of quantitative lung nodule volume change measurement in malignant and benign lung nodules. The coefficient of variation of the measurement was found to be tight even for small nodules <300 $mm^3$ demonstrating the accuracy of volumetric measurement when analyzing CT images obtained with a high-quality CT image acquisition protocol (Table 3). Such measurement accuracy may be clinically important as semi-automated volumetric nodule measurements of CT scans acquired with high quality image acquisition (with a time interval between CT scans ≥ 120 days) resulted in a sensitivity of 89 % and specificity of 100 % compared to 75% sensitivity

and 98% specificity resulting from 2-D measurements using the same set of scans. This suggests that use of CT slice thicknesses <= 1.0 mm with appropriate follow-up time interval may improve the performance of semi-automated volumetric assessment and warrants further study by the global lung cancer community.

Other potential benefits arising from the development of ELIC include the use of a globally distributed structured lung cancer screening database. ELIC requires all participating sites to provide metadata using a common nomenclature and data structure. All ELIC sites providing screening data in this report used the open source VAPALS-ELCAP data dictionary. The data assembled for this manuscript therefore comprised a pilot effort with data standardization across the 7 ELIC contributing sites. This approach also could support global lung cancer imaging research efforts moving forward. This would be a significant contribution of IASLC to the refinement of future lung cancer screening management. Finally, ELIC might also be used to establish a curated IASLC image research collection for both training and evaluation. An archive of screening cases assembled with appropriate permissions utilizing the thoracic CT images and associated clinical ground truth data may be of considerable value as a global IASLC resource.

Lessons learned from this pilot study include the issues associated with the variability of the submitted image and clinical data acquired across the globe. Despite having provided contributing sites with a specific acquisition protocol, major problems occurred with automated analysis by AI and quantitative algorithms. **Figure 2** shows that DICOM CT image data issues and nodule location issues significantly impacted the number of high-quality cases for volumetric analysis in ELIC. In the future we plan to improve the CT image quality validation methods and tools in order to better ensure that only data that adhere to high levels of CT image quality can be entered e.g. verification of CT slice thickness, spacing and continuity. ELIC is also working with international lung cancer screening sites to contribute more cases to ELIC to expand this valuable open resource to facilitate AI research in lung screening.

**Limitations**

This pilot study has potential limitations in terms of reproducibility and generalizability given the study size and preliminary nature of the reported findings. In creating ELIC as an open-source image resource with donated images and meta data, we are providing resources so independent verification of our reported finding is readily doable. We encourage other researchers to critically assess the findings of this report and conduct additional research to improve the process of thoracic CT screening.

**CONCLUSION**

ELIC is now a functional resource to advance deep learning AI and quantitative imaging research studies on a diverse, globally distributed CT screening dataset. Expansion of ELIC can provide a major global resource for lung cancer imaging research.

Acknowledgement:

We thank all the screening participants that contributed CT images and associated clinical data to enable the evaluations described in this manuscript and all the institutions that supported the development of this open imaging research environment by allowing case contributions to this resource. We thank the IASLC and Astra Zeneca for the financial support to enable this research report. We thank the many research nurses, data managers and informaticians whose efforts allowed us to populate ELIC.

**REFERENCES**


1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021; 71: 209– 49.

2. Canadian Cancer Statistics Advisory Committee in collaboration with the Canadian Cancer Society T, ON: Canadian Cancer Society;. . Canadian Cancer Statistics 2021. : Statistics Canada and the Public Health Agency of Canada; 2021.

3. Henschke CI, McCauley DI, Yankelevitz DF, Naidich DP, McGuinness G, Miettinen OS, Libby DM, Pasmantier MW, Koizumi J, Altorki NK, Smith JP. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999; 354:99-105 PMID:10408484.

4. 4 Henschke CI, Naidich DP, Yankelevitz DF, McGuinness G, McCauley DI, Smith JP, Libby D, Pasmantier M, Vazquez M, Koizumi J, Flieder D, Altorki N, Miettinen OS. Early Lung Cancer Action Project: initial findings on repeat screenings. Cancer 2001; 92:153-9 PMID:11443621.

5. Aberle, D.R., et al., Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med, 2011. 365(5): p. 395-409.

6. de Koning, H.J., et al., Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. N Engl J Med, 2020. 382(6): p. 503-513.

7. Pastorino, U., et al., Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. Ann Oncol, 2019. 30(7): p. 1162-1169.

8. Field, J.K., et al., Lung cancer mortality reduction by LDCT screening: UKLS randomised trial results and international meta-analysis. Lancet Reg Health Eur, 2021. 10: p. 100179.

9. Bonney, A., et al., Impact of low-dose computed tomography (LDCT) screening on lung cancer-related mortality. Cochrane Database Syst Rev, 2022. 8(8): p. Cd013829.

10. Network, L.C.P. Interactive map of lung cancer screening 2022; Available from: https://www.lungcancerpolicynetwork.com/interactive-map-of-lung-cancer-screening/.

11. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G, Naidich DP, Shetty S. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. 2019

Jun;25(6):954-961. doi: 10.1038/s41591-019-0447-x. Epub 2019 May 20. Erratum in: Nat Med. 2019 Aug;25(8):1319. PMID: 31110349.

12. Zhang C, Sun X, Dang K, Li K, Guo XW, Chang J, Yu ZQ, Huang FY, Wu YS, Liang Z, Liu ZY, Zhang XG, Gao XL, Huang SH, Qin J, Feng WN, Zhou T, Zhang YB, Fang WJ, Zhao MF, Yang XN, Zhou Q, Wu YL, Zhong WZ. Toward an Expert Level of Lung Cancer Detection and Classification Using a Deep Convolutional Neural Network. Oncologist. 2019 Sep;24(9):1159-1165. doi: 10.1634/theoncologist.2018-0908. Epub 2019 Apr 17. PMID: 30996009; PMCID: PMC6738288.

13. Min Y, Hu L, Wei L, Nie SD. Computer-aided detection of pulmonary nodules based on convolutional neural networks: a review. Phys Med Biol. 2022 Feb 18. doi: 10.1088/1361-6560/ac568e. Epub ahead of print. PMID: 35180709.

14. Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B, Kostis WJ. On measuring the change in size of pulmonary nodules. IEEE Trans Med Imaging. 2006 Apr;25(4):435-50. doi: 10.1109/TMI.2006.871548. PMID: 16608059.

15. Das M, Ley-Zaporozhan J, Gietema HA, Czech A, Mühlenbruch G, Mahnken AH, Katoh M, Bakai A, Salganicoff M, Diederich S, Prokop M, Kauczor HU, Günther RW, Wildberger JE. Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners. Eur Radiol. 2007 Aug;17(8):1979-84. doi: 10.1007/s00330-006-0562-1. Epub 2007 Jan 6. PMID: 17206420.

16. Mikhael PG, Wohlwend J, Yala A, Karstens L, et al. Sybil: A Validated Deep Learning Model to Predict Future Lung Cancer Risk From a Single Low-Dose Chest Computed Tomography . J Clin Oncol. 2023 Jan 12:JCO2201345. doi: 10.1200/JCO.22.01345. Online ahead of print.

17. Mulshine JL, Avila RS, Conley E, Devaraj A, Ambrose LF, Flanagan T, Henschke CI, Hirsch FR, Janz R, Kakinuma R, Lam S, McWilliams A, Van Ooijen PMA, Oudkerk M, Pastorino U, Reeves A, Rogalla P, Schmidt H, Sullivan DC, Wind HHJ, Wu N, Wynes M, Xueqian X, Yankelevitz DF, Field JK. The International Association for the Study of Lung Cancer Early Lung Imaging Confederation. JCO Clin Cancer Inform. 2020 Feb;4:89-99. doi: 10.1200/CCI.19.00099. PMID: 32027538; PMCID: PMC7053806.

18. Hallock H, Marshall SE, 't Hoen PAC, Nygård JF, Hoorne B, Fox C, Alagaratnam S. Federated Networks for Distributed Analysis of Health Data. Front Public Health. 2021 Sep 30;9:712569. doi: 10.3389/fpubh.2021.712569. PMID: 34660512; PMCID: PMC8514765.

19. Scherer J, Nolden M, Kleesiek J, Metzger J, Kades K, Schneider V, Bach M, Sedlaczek O, Bucher AM, Vogl TJ, Grünwald F, Kühn JP, Hoffmann RT, Kotzerke J, Bethge O, Schimmöller L, Antoch G, Müller HW, Daul A, Nikolaou K, la Fougère C, Kunz WG,

Ingrisch M, Schachtner B, Ricke J, Bartenstein P, Nensa F, Radbruch A, Umutlu L, Forsting M, Seifert R, Herrmann K, Mayer P, Kauczor HU, Penzkofer T, Hamm B, Brenner W, Kloeckner R, Düber C, Schreckenberger M, Braren R, Kaissis G, Makowski M, Eiber M, Gafita A, Trager R, Weber WA, Neubauer J, Reisert M, Bock M, Bamberg F, Hennig J, Meyer PT, Ruf J, Haberkorn U, Schoenberg SO, Kuder T, Neher P, Floca R, Schlemmer HP, Maier-Hein K. Joint Imaging Platform for Federated Clinical Data Analytics. JCO Clin Cancer Inform. 2020 Nov;4:1027-1038. doi: 10.1200/CCI.20.00045. PMID: 33166197; PMCID: PMC7713526.

20. Erickson BJ, Fajnwaks P, Langer SG, Perry J. Multisite Image Data Collection and Management Using the RSNA Image Sharing Network. Transl Oncol. 2014 Feb 1;7(1):36-9. doi: 10.1593/tlo.13799. PMID: 24772205; PMCID: PMC3998697.

21. Lewis JA, Spalluto LB, Henschke CI, Yankelevitz DF, Aguayo SM, Morales P, Avila R, Audet CM, Prusaczyk B, Lindsell CJ, Callaway-Lane C, Dittus RS, Vogus TJ, Massion PP, Limper HM, Kripalani S, Moghanaki D, Roumie CL. Protocol to evaluate an enterprise-wide initiative to increase access to lung cancer screening in the Veterans Health Administration. Clin Imaging. 2021 May;73:151-161. doi: 10.1016/j.clinimag.2020.11.059. Epub 2020 Dec 26. PMID: 33422974; PMCID: PMC8479827.

22. de Torres Juan P., Bastarrika, G, Wisnivesky JP et al. Assessing the Relationship Between Lung Cancer Risk and Emphysema Detected on Low-Dose CT of the Chest. Chest 2007;132 (6):1932-1938.

23. Stolk J, Putter H, Bakker EM, et al. Progression parameters for emphysema: a clinical investigation. Respir Med. 2007;101(9):1924-1930.

24. Barbosa EM Jr, Song G, Tustison N, Kreider M, Gee JC, Gefter WB, Torigian DA. Computational analysis of thoracic multidetector row HRCT for segmentation and quantification of small airway air trapping and emphysema in obstructive pulmonary disease. Acad Radiol. 2011 Oct;18(10):1258-69. doi: 10.1016/j.acra.2011.06.004. PMID: 21893294. Barbosa EM Jr, Song G, Tustison N, Kreider M, Gee JC, Gefter WB, Torigian DA. Computational analysis of thoracic multidetector row HRCT for segmentation and quantification of small airway air trapping and emphysema in obstructive pulmonary disease. Acad Radiol. 2011 Oct;18(10):1258-69. doi: 10.1016/j.acra.2011.06.004. PMID: 21893294.

25. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Advances in Neural Information Processing Systems 32 [Internet]. Curran Associates, Inc.; 2019. p. 8024–35.

26. Isensee F, Jaeger PF, Kohl SAA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18:203–211.

27. Krishnan K, Ibanez L, Turner WD, Jomier J, Avila RS. An open-source toolkit for the volumetric measurement of CT lung lesions. Opt Express. 2010 Jul 5;18(14):15256-66. doi: 10.1364/OE.18.015256. PMID: 20640012.

28. Mulshine JL, Gierada DS, Armato SG 3rd, Avila RS, Yankelevitz DF, Kazerooni EA, McNitt-Gray MF, Buckler AJ, Sullivan DC. Role of the Quantitative Imaging Biomarker Alliance in optimizing CT for the evaluation of lung cancer screen-detected nodules. J Am Coll Radiol. 2015 Apr;12(4):390-5. doi: 10.1016/j.jacr.2014.12.003. PMID: 25842017.

29. Rajendran S, Obeid JS, Binol H, D Agostino R Jr, Foley K, Zhang W, Austin P, Brakefield J, Gurcan MN, Topaloglu U. Cloud-Based Federated Learning Implementation Across Medical Centers. JCO Clin Cancer Inform. 2021 Jan;5:1-11. doi: 10.1200/CCI.20.00060. PMID: 33411624; PMCID: PMC8140794.

30. Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F. Federated Learning for Healthcare Informatics. J Healthc Inform Res. 2021;5(1):1-19. doi: 10.1007/s41666-020-00082-4. Epub 2020 Nov 12. PMID: 33204939; PMCID: PMC7659898.

31. Gallardo-Estrella L, Pompe E, de Jong PA, Jacobs C, van Rikxoort EM, Prokop M, Sánchez CI, van Ginneken B. Normalized emphysema scores on low dose CT: Validation as an imaging biomarker for mortality. PLoS One. 2017 Dec 11;12(12):e0188902. doi: 10.1371/journal.pone.0188902. PMID: 29227997; PMCID: PMC5724850.

32. Mascalchi M, Camiciottoli G, Diciotti S. Lung densitometry: why, how and when. J Thorac Dis. 2017 Sep;9(9):3319-3345. doi: 10.21037/jtd.2017.08.17. PMID: 29221318; PMCID: PMC5708390.

**Table 1:** Lung cancer screening participant demographics for the ELIC dataset.

| | **No Nodules** | **Benign Nodules** | **Lung Cancer** |
|---|---|---|---|
| **Participant Cases** | 172 | 402 | 123 |
| **Nodule Count** | 0 | 1383 | 475 |
| **Mean Age (+- SD)** | 61.70 ± 7.62 | 62.58 ± 8.48 | 65.15 ± 8.43 |
| **Gender = Male** | 57.56 % | 52.99 % | 56.10 % |
| **Smoking Status** | | | |
| - **Past** | 34.88 % | 42.54 % | 41.46 % |
| - **Current** | 48.84 % | 41.04 % | 47.15 % |
| - **Never** | 16.28 % | 16.42 % | 11.38 % |
| **Mean Pack Years (+- SD)** | 34.52 ± 24.56 | 38.60 ± 28.64 | 41.71 ± 28.71 |

**Table 2**: A listing of spokes from ordered according to the CT slice thickness used to scan screening participants. Emphysema metric ratios were evaluated comparing lung cancer cases to benign cases showing that LAV950 applied to thick CT image slices had the highest emphysema metric ratios (red text) between lung cancer cases and non-lung cancer cases.

| Spoke | Slice Thickness | Lung Volume | Mean HU | Median HU | Std Dev HU | Perc15 | LAV950 % |
|---|---|---|---|---|---|---|---|
| **1** | 0.75 mm | 0.94 | 1.01 | 1.01 | 0.96 | 1.00 | 1.04 |
| **2** | 1.00 mm | 0.98 | 1.01 | 1.01 | 0.98 | 1.00 | 1.02 |
| **3** | 1.00 mm | 1.08 | 1.00 | 1.01 | 0.96 | 1.00 | 0.98 |
| **4** | 1.25 mm | 0.92 | 0.98 | 0.99 | 0.93 | 1.00 | 0.99 |
| **5** | 1.25 mm | 1.02 | 0.98 | 0.99 | 0.96 | 0.98 | 1.21 |
| **6** | 2.00 mm | 1.29 | 1.04 | 1.03 | 1.05 | 1.02 | 1.36 |
| **7** | 3.00 mm | 1.03 | 1.02 | 1.01 | 1.05 | 1.01 | 1.34 |

**Table 3**: The mean lung nodule volume change and the Coefficient of Variation (COV) for malignant and benign solid nodules in data subset 2.

| | Nodule Volume $mm^3$ | Mean Volume Change (%) | COV |
|---|---|---|---|
| **Non-Lung Cancer Cases** | < 300 | 6.6 | 11.0 |
| | >= 300 | -101.4 | -4.7 |
| **Lung Cancer Cases** | < 300 | 346.9 | 0.9 |
| | >= 300 | 382.5 | 1.1 |

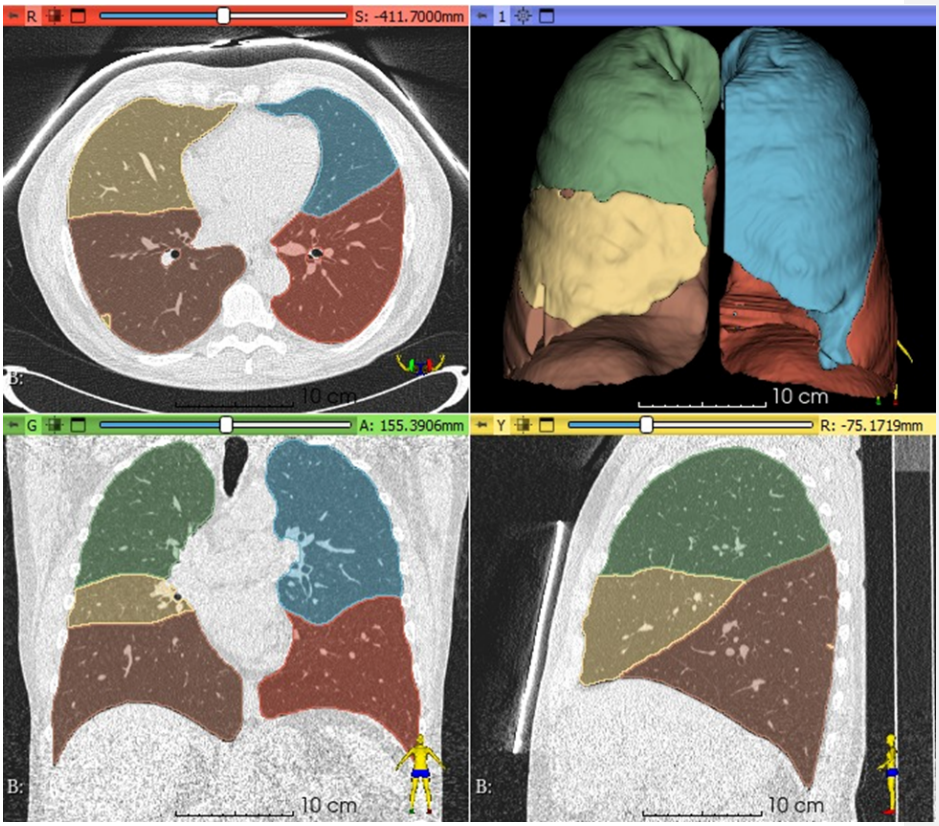


**Figure 1:** Fully automated deep learning lung and lobe segmentation of a low dose CT lung cancer screening scan.

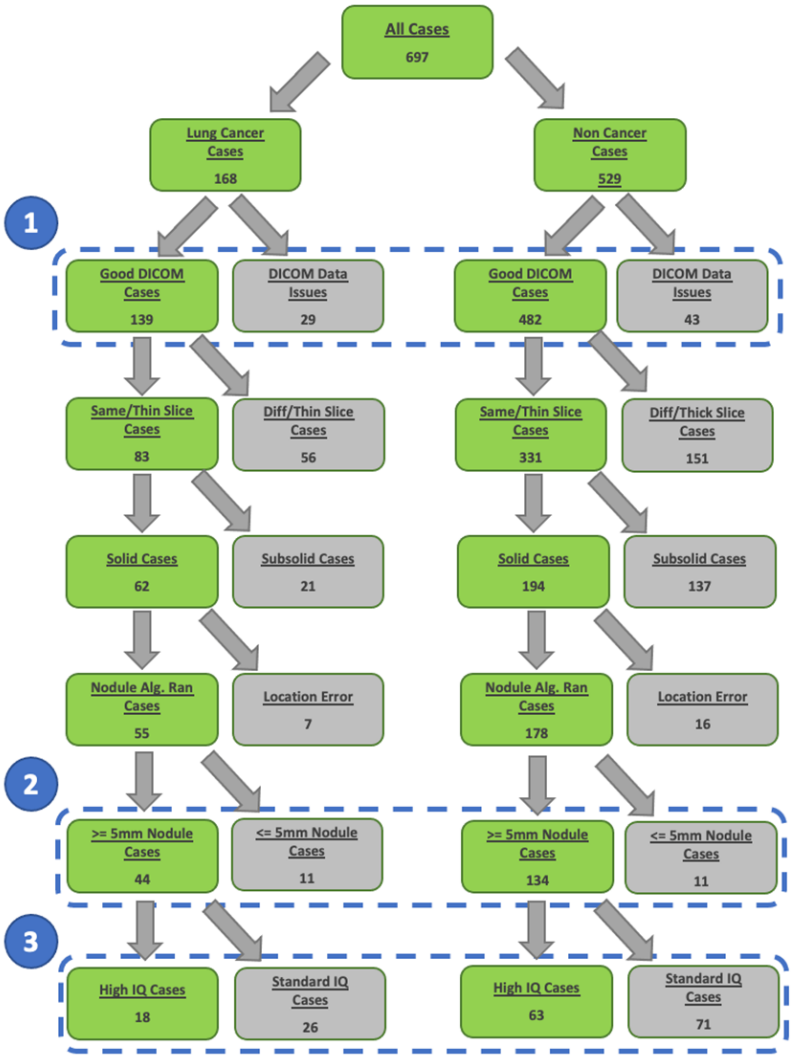


**Figure 2:** Construction of three data subsets for (1) the evaluation of a deep learning AI segmentation and quantitative emphysema algorithm, and (2) the evaluation of a small solid lung nodule volume change algorithm on solid lung nodules, and (3) the evaluation of the same small solid lung nodule volume change algorithm on high CT image quality data.

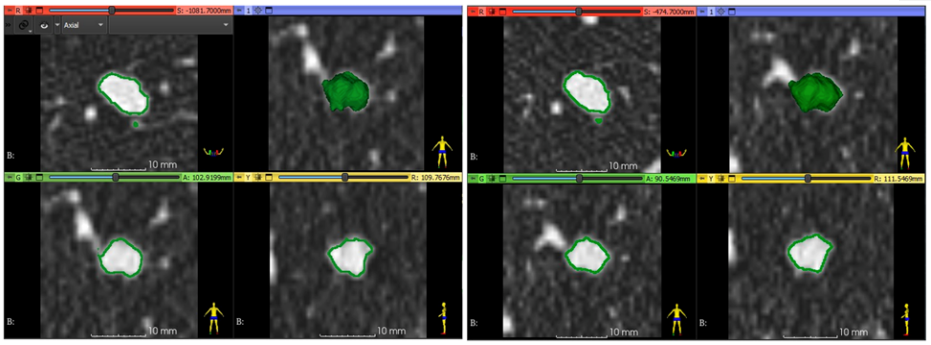

**Figure 3:** Semi-automated volumetric measurement of a benign solid lung nodule at time point 1 (left) and time point 2 (right) showing a small increase in volumetric size. Each time point image panel shows axial (top-left), coronal with 3D nodule surface shown in green (top-right), coronal (bottom-left), and sagittal (bottom-right) image reformats.

**Commented [JM1]:** Labeling quadrants with number or letters used in the legend may simplify the navigation for the reader

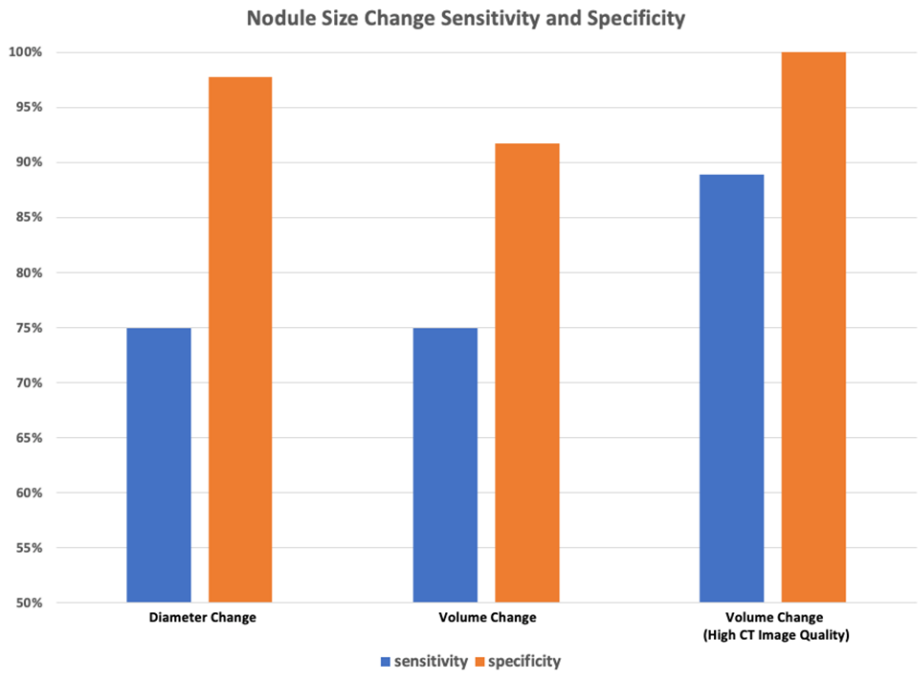


Figure 4: Performance of radiologist diameter (left), semi-automated volume change (middle), and semi-automated volume change using only high-quality CT image data and segmentation results (right) at classifying lung cancer malignancy.