National Lung Screening Trial (NLST)

Redirection Notice

This page will redirect to https://doi.org/10.7937/TCIA.HMQ8-J677 in about 5 seconds.

Summary

Background: The aggressive and heterogeneous nature of lung cancer has thwarted efforts to reduce mortality from this cancer through the use of screening. The advent of low-dose helical computed tomography (CT) altered the landscape of lung-cancer screening, with studies indicating that low-dose CT detects many tumors at early stages. The National Lung Screening Trial (NLST) was conducted to determine whether screening with low-dose CT could reduce mortality from lung cancer.

Methods: From August 2002 through April 2004, we enrolled 53,454 persons at high risk for lung cancer at 33 U.S. medical centers. Participants were randomly assigned to undergo three annual screenings with either low-dose CT (26,722 participants) or single-view posteroanterior chest radiography (26,732). Data were collected on cases of lung cancer and deaths from lung cancer that occurred through December 31, 2009. This dataset includes the low-dose CT scans from 26,254 of these subjects, as well as digitized histopathology images from 451 subjects.

Results: The rate of adherence to screening was more than 90%. The rate of positive screening tests was 24.2% with low-dose CT and 6.9% with radiography over all three rounds. A total of 96.4% of the positive screening results in the low-dose CT group and 94.5% in the radiography group were false positive results. The incidence of lung cancer was 645 cases per 100,000 person-years (1060 cancers) in the low-dose CT group, as compared with 572 cases per 100,000 person-years (941 cancers) in the radiography group (rate ratio, 1.13; 95% confidence interval [CI], 1.03 to 1.23). There were 247 deaths from lung cancer per 100,000 person-years in the low-dose CT group and 309 deaths per 100,000 person-years in the radiography group, representing a relative reduction in mortality from lung cancer with low-dose CT screening of 20.0% (95% CI, 6.8 to 26.7; P=0.004). The rate of death from any cause was reduced in the low-dose CT group, as compared with the radiography group, by 6.7% (95% CI, 1.2 to 13.6; P=0.02).

Conclusions: Screening with the use of low-dose CT reduces mortality from lung cancer. (Funded by the National Cancer Institute; National Lung Screening Trial ClinicalTrials.gov number, NCT00047385).

Data Availability: A summary of the National Lung Screening Trial and its available datasets are provided on the Cancer Data Access System (CDAS). CDAS is maintained by Information Management System (IMS), contracted by the National Cancer Institute (NCI) as keepers and statistical analyzers of the NLST trial data. The full clinical data set from NLST is available through CDAS. Users of TCIA can download without restriction a publicly distributable subset of that clinical data, along with the CT and Histopathology images collected during the trial. (These previously were restricted.)

Data Access

Data Access

Data Type	Download all or Query/Filter	License
Radiology CT Images (26254 subjects, DICOM, 11.3 TB)	*Download Search* This link downloads the entire collection, which is quite large. See the Detailed Description tab for options to download the collection in smaller chunks. (Download requires the NBIA Data Retriever)	CC BY 4.0
Tissue Slide Images - Primary Tumor (451 subjects, 1225 files, SVS, 775 GB)	*Download Search* ***Additional images are available***: See Detailed Description. (Download and apply the IBM-Aspera-Connect plugin to your browser to retrieve this faspex package)	CC BY 4.0
Clinical data including data dictionaries (SAS, ZIP, 25 MB)	*Download* Provided in SAS format in one compressed file (.zip); includes data and dictionaries.	CC BY 4.0

Additional histopathology slide images Table 1 for which the participants have no Baseline Questionnaire data (2 subjects, DOCX, 13 KB)	*Download*	CC BY 4.0
Histopathology additional slide images for which the participants have no Baseline Questionnaire data (2 subjects, 4 files, SVS)	*Download* (Download and apply the IBM-Aspera-Connect plugin to your browser)	CC BY 4.0
Additional histopathology slide images Table 2 for participants with Second Primary Tumors as well as those included in the "standard" package (10 subjects, 23 images, DOCX, 23 KB)	*Download*	CC BY 4.0
Histopathology additional slide images for participants with Second Primary Tumors as well as those included in the "standard" package (10 subjects, 23 files, SVS, 18.7 GB)	*Download* (Download and apply the IBM-Aspera-Connect plugin to your browser)	CC BY 4.0

Additional Resources for this Dataset

The NCI Cancer Research Data Commons (CRDC) provides access to additional data and a cloud-based data science infrastructure that connects data sets with analytics tools to allow users to share, integrate, analyze, and visualize cancer research data.

- Imaging Data Commons (IDC) (Imaging Data)

The following external resources have been made available by the data submitters. These are not hosted or supported by TCIA, but may be useful to the researchers utilizing this collection

- Clinical data

This is a subset of the full clinical data. If you need the full clinical data, please visit the Cancer Data Access System (CDAS) system.
Detailed Description

Detailed Description

Collection Statistics	Radiology	Pathology
Modalities	CT	Aperio
Number of Patients	26,254	451
Number of Studies	73,118	
Number of Series	203,099	
Number of Images	21,082,502	1,225 (optionally + 4 + 23)
Images Size (TB)	11.3 TB	775 GB

The full CT data (manifest-NLST_allCT.tcia) occupy 11.3 terabytes when downloaded. For convenience, you can "Search" to access all the files, or you can download in chunks.

The pathology slide data:

1. Primary Tumor slides (faspex) Primary Tumor slides (the standard package), 1225 files.
2. Additional slides (faspex) Additional histopathology slide images for which the participants have no Baseline Questionnaire data (4 slides) Detail in Table 1.
3. Second Primary-Tumor slides (faspex) Additional histopathology slide images for participants with Second Primary Tumors as well as those included in the "standard" package (23 slides) Detail in Table 2.

NLST Design & Process, Protocol Documents, and Results: https://cdas.cancer.gov/learn/nlst/main-findings/

- Overview, study design, recruitment methods, endpoint verification process (determining cause of death), quality control (LSS and ACRIN).
- LSS Manual of Operations (MOOP) and ACRIN Protocol.
- DSMB announcement of results (10/28/2010).
- Feasibility Study for the NLST, Psychosocial and Behavioral Issues, Technical Publications. https://cdas.cancer.gov/learn/nlst/main-findings/
- Browse Publications : https://cdas.cancer.gov/publications/?study=nlst
- Browse projects : https://cdas.cancer.gov/approved-projects/?data_types=nlst_data

NLST Data Collected: https://biometry.nci.nih.gov/cdas/learn/nlst/data-collected/

- Questionnaires, screening, diagnostic procedures, cancer diagnosis, treatment, progression, mortality, contamination.

Biospecimens Collected

Formalin-fixed paraffin embedded (FFPE) tissue specimens are available for a subset of the NLST participants who developed lung cancer during the trial. Donor blocks were obtained from local pathology laboratories and tissue cores (0.6mm) were extracted from them to construct tissue microarrays (TMA). Tissue cores were sampled from primary main invasive tumor histology, secondary tumor histology, carcinoma in situ, adjacent normal lung tissue, metastatic lesion from lymph node(s) and/or distant sites, benign (un-involved) lymph node, proximal and/or distal bronchi.

In total, tissue materials were collected from 438 lung cancer cases. All have cores arrayed across nine TMAs, one of which only contains tissue collected after neoadjuvant treatment. 434 of these also have loose cores available for nucleic acid extraction. On average, each TMA contains 504 cores from 48 subjects.

Applications for access to these specimens can be submitted under the PLCO Etiologic and Early Marker Studies Program (EEMS). The application review process opens twice a year, once in the winter and once in the summer. For more information about EEMS and to initiate an application visit the PLCO EEMS Application page. When filling out the application, specify “NLST Tissue” under the case definition.

Citations & Data Usage Policy

Citations & Data Usage Policy

Users must abide by the TCIA Data Usage Policy and Restrictions. Attribution should include references to the following citations:

Data Citation

National Lung Screening Trial Research Team. (2013). **Data from the National Lung Screening Trial (NLST) [Data set].** The Cancer Imaging Archive. https://doi.org/10.7937/TCIA.HMQ8-J677

Publication Citation

National Lung Screening Trial Research Team; Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD (2011). **Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening.** New England Journal of Medicine, 365(5), 395–409. https://doi.org/10.1056/nejmoa1102873

TCIA Citation

Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. **The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository**, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. DOI: https://doi.org/10.1007/s10278-013-9622-7

Other Publications Using This Data

The Collection authors suggest the below will give context to this dataset:

- National Lung Screening Trial Research Team. (2011). **The national lung screening trial: overview and study design.** Radiology, 258(1), 243-253. doi:https://doi.org/10.1148/radiol.10091808
- Pinsky, P. F., Gierada, D. S., Nath, H., Kazerooni, E. A., & Amorosa, J. (2013). ROC curves for low-dose CT in the National Lung Screening Trial. J Med Screen, 20(3), 165-168. doi:10.1177/0969141313500666
- Pinsky, P. F., Gierada, D. S., Nath, P. H., Kazerooni, E., & Amorosa, J. (2013). National lung screening trial: variability in nodule detection rates in chest CT studies. Radiology, 268(3), 865-873. doi:10.1148/radiol.13121530
- National Lung Screening Trial Research Team, Aberle, D. R., Adams, A. M., Berg, C. D., Black, W. C., Clapp, J. D., . . . Sicks, J. D. (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. New England Journal of Medicine, 365(5), 395-409. doi: https://doi.org/10.1056/NEJMoa1102873

TCIA maintains a list of publications which leverage TCIA data. If you have a manuscript you'd like to add please contact TCIA's Helpdesk.

Note: IMS/CDAS maintains a separate list of publications related to NLST data: https://cdas.cancer.gov/publications/?study=nlst

1. Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., . . . Shetty, S. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. doi:10.1038/s41591-019-0447-x
2. Balagurunathan, Y., Schabath, M. B., Wang, H., Liu, Y., & Gillies, R. J. (2019). Quantitative Imaging features Improve Discrimination of Malignancy in Pulmonary nodules. Sci Rep, 9(1). doi:10.1038/s41598-019-44562-z
3. Bartel, S. T., Bierhals, A. J., Pilgram, T. K., Hong, C., Schechtman, K. B., Conradi, S. H., & Gierada, D. S. (2011). Equating quantitative emphysema measurements on different CT image reconstructions. Medical Physics, 38(8), 4894-4902. doi:10.1118/1.3615624
4. Cherezov, D., Goldgof, D., Hall, L., Gillies, R., Schabath, M., Müller, H., & Depeursinge, A. (2019). Revealing Tumor Habitats from Texture Heterogeneity Analysis for Classification of Lung Cancer Malignancy and Aggressiveness. Sci Rep, 9(1), 4500. doi:10.1038/s41598-019-38831-0
5. Church, T. R., Black, W. C., Aberle, D. R., Berg, C. D., Clingan, K. L., Duan, F., . . . Baum, S. (2013). Results of initial low-dose computed tomographic screening for lung cancer. N Engl J Med, 368(21), 1980-1991. doi:10.1056/NEJMoa1209120

6. Foley, F., Rajagopalan, S., Raghunath, S. M., Boland, J. M., Karwoski, R. A., Maldonado, F., . . . Peikert, T. (2016). Computer-aided nodule assessment and risk yield risk management of adenocarcinoma: the future of imaging? Paper presented at the Seminars in thoracic and cardiovascular surgery.
7. Gierada, D. S., Guniganti, P., Newman, B. J., Dransfield, M. T., Kvale, P. A., Lynch, D. A., & Pilgram, T. K. (2011). Quantitative CT assessment of emphysema and airways in relation to lung cancer risk. Radiology, 261(3), 950-959. doi:https://doi.org/10.1148/radiol.11110542
8. Gierada, D. S., Pinsky, P., Nath, H., Chiles, C., Duan, F., & Aberle, D. R. (2014). Projected outcomes using different nodule sizes to define a positive CT lung cancer screening examination. Journal of the National Cancer Institute, 106(11), dju284. doi:10.1093/jnci/dju284
9. Gunawan, R., Tran, Y., Zheng, J., Nguyen, H., & Chai, R. (2022). Image Recovery from Synthetic Noise Artifacts in CT Scans Using Modified U-Net. Sensors (Basel), 22(18). doi:https://doi.org/10.3390/s22187031
10. Jamdade, V. A. (2022). Explainable Lung Nodule Malignancy Classification from CT Scans. (M.S. Thesis). University of Maryland, Baltimore County, USA, University of Maryland, Baltimore CountyProQuest Dissertations Publishing. Retrieved from https://dissexpress.proquest.com/dxweb/results.html?QryTxt=&pubnum=29997250
11. Jeon, K. N., Goo, J. M., Lee, C. H., Lee, Y., Choo, J. Y., Lee, N. K., . . . Gierada, D. S. (2012). Computer-aided nodule detection and volumetry to reduce variability between radiologists in the interpretation of lung nodules at low-dose screening CT. Investigative radiology, 47(8), 457. doi: 10.1097/RLI.0b013e318250a5aa
12. Lo, S. B., Freedman, M. T., Gillis, L. B., White, C. S., & Mun, S. K. (2018). JOURNAL CLUB: Computer-Aided Detection of Lung Nodules on CT With a Computerized Pulmonary Vessel Suppressed Function. American Journal of Roentgenology, 210(3), 480-488. doi:10.2214/AJR.17.18718
13. Mikhael, P. G., Wohlwend, J., Yala, A., Karstens, L., Xiang, J., Takigami, A. K., . . . Barzilay, R. (2023). Sybil: A Validated Deep Learning Model to Predict Future Lung Cancer Risk From a Single Low-Dose Chest Computed Tomography. J Clin Oncol, JCO2201345. doi:https://doi.org/10.1200/JCO.22.01345
14. Patz Jr, E. F., Greco, E., Gatsonis, C., Pinsky, P., Kramer, B. S., & Aberle, D. R. (2016). Lung cancer incidence and mortality in National Lung Screening Trial participants who underwent low-dose CT prevalence screening: a retrospective cohort analysis of a randomised, multicentre, diagnostic screening trial. The Lancet Oncology, 17(5), 590-599. doi:https://doi.org/10.1016/S1470-2045(15)00621-X
15. Perez-Morales, J., Tunali, I., Stringfield, O., Eschrich, S. A., Balagurunathan, Y., Gillies, R. J., & Schabath, M. B. (2020). Peritumoral and intratumoral radiomic features predict survival outcomes among patients diagnosed in lung cancer screening. Sci Rep, 10(1), 10528. doi:https://doi.org/10.1038/s41598-020-67378-8
16. Petousis, P., Han, S. X., Aberle, D., & Bui, A. A. (2016). Prediction of lung cancer incidence on the low-dose computed tomography arm of the National Lung Screening Trial: A dynamic Bayesian network. Artificial intelligence in medicine, 72, 42-55. doi:https://doi.org/10.1016/j.artmed.2016.07.001
17. Pilgram, T. K., Quirk, J. D., Bierhals, A. J., Yusen, R. D., Lefrak, S. S., Cooper, J. D., & Gierada, D. S. (2010). Accuracy of emphysema quantification performed with reduced numbers of CT sections. American Journal of Roentgenology, 194(3), 585-591. doi:10.2214/AJR.09.2709
18. Pinsky, P. F., Nath, P. H., Gierada, D. S., Sonavane, S., & Szabo, E. (2014). Short-and long-term lung cancer risk associated with noncalcified nodules observed on low-dose CT. Cancer prevention research, 7(12), 1179-1185. doi:10.1158/1940-6207.CAPR-13-0438
19. Pu, L., Gezer, N. S., Ashraf, S. F., Ocak, I., Dresser, D. E., & Dhupar, R. (2022). Automated segmentation of five different body tissues on computed tomography using deep learning. Med Phys. doi:https://doi.org/10.1002/mp.15932
20. Reeves, A. P., Xie, Y., & Jirapatnakul, A. (2016). Automated pulmonary nodule CT image characterization in lung cancer screening. International Journal of Computer Assisted Radiology and Surgery, 11(1), 73-88. doi: 10.1007/s11548-015-1245-7
21. Salama, W. M., Aly, M. H., & Elbagoury, A. M. (2021). Lung Images Segmentation and Classification Based on Deep Learning: A New Automated CNN Approach. Journal of Physics: Conference Series, 2128(1). doi:10.1088/1742-6596/2128/1/012011
22. Schreuder, A., Jacobs, C., Gallardo-Estrella, L., Prokop, M., Schaefer-Prokop, C. M., & van Ginneken, B. (2019). Predicting all-cause and lung cancer mortality using emphysema score progression rate between baseline and follow-up chest CT images: A comparison of risk model performances. PLoS One, 14(2), e0212756. doi:https://doi.org/10.1371/journal.pone.0212756
23. Schreuder, A., van Ginneken, B., Scholten, E. T., Jacobs, C., Prokop, M., Sverzellati, N., . . . Schaefer-Prokop, C. M. (2018). Classification of CT pulmonary opacities as perifissural nodules: reader variability. Radiology, 288(3), 867-875. doi:https://doi.org/10.1148/radiol.2018172771
24. Shields, B., & Ramachandran, P. (2023). Generating missing patient anatomy from partially acquired cone-beam computed tomography images using deep learning: a proof of concept. Phys Eng Sci Med. doi:https://doi.org/10.1007/s13246-023-01302-y
25. Singh, S., Gierada, D. S., Pinsky, P., Sanders, C., Fineberg, N., Sun, Y., . . . Nath, H. (2012). Reader variability in identifying pulmonary nodules on chest radiographs from the national lung screening trial. Journal of thoracic imaging, 27(4), 249. doi:10.1097/RTI.0b013e318256951e
26. Singh, S., Pinsky, P., Fineberg, N. S., Gierada, D. S., Garg, K., Sun, Y., & Nath, P. H. (2011). Evaluation of reader variability in the interpretation of follow-up CT scans at lung cancer screening. Radiology, 259(1), 263-270. doi:10.1148/radiol.10101254
27. Torres, F. S., Akbar, S., Raman, S., Yasufuku, K., Schmidt, C., Hosny, A., . . . Leighl, N. B. (2021). End-to-End Non-Small-Cell Lung Cancer Prognostication Using Deep Learning Applied to Pretreatment Computed Tomography. JCO Clin Cancer Inform, 5, 1141-1150. doi:10.1200/cci.21.00096
28. Uthoff, J. M. (2019). Cancer Risk Assessment Using Quantitative Imaging Features from Solid Tumors and Surrounding Structures. (Ph.D. Dissertation). The University of Iowa, Ann Arbor, United States. Retrieved from https://www.proquest.com/dissertations-theses/cancer-risk-assessment-using-quantitative-imaging/docview/2306303717/se-2?accountid=142023 (2306303717, 13858412)
29. Wu, D., Liu, R., Levitt, B., Riley, T., & Baumgartner, K. (2016). Evaluating long-term outcomes via computed tomography in lung cancer screening. J Biom Biostat, 7(313), 2. doi:10.4172/2155-6180.1000313
30. Yip, R., Henschke, C. I., Xu, D. M., Li, K., Jirapatnakul, A., & Yankelevitz, D. F. (2017). Lung Cancers Manifesting as Part-Solid Nodules in the National Lung Screening Trial. American Journal of Roentgenology, 208(5), 1011-1021. doi:10.2214/Ajr.16.16930
31. Yip, R., Yankelevitz, D. F., Hu, M., Li, K., Xu, D. M., Jirapatnakul, A., & Henschke, C. I. (2016). Lung cancer deaths in the National Lung Screening Trial attributed to nonsolid nodules. Radiology, 281(2), 589-596. doi:https://doi.org/10.1148/radiol.2016152333
32. Zhao, T., & Yin, Z. (2021). Airway Anomaly Detection by Prototype-Based Graph Neural Network. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention, Strasbourg, France.
33. Zhu, C. S., Pinsky, P. F., Moler, J. E., Kukwa, A., Mabie, J., Rathmell, J. M., . . . Berg, C. D. (2017). Data sharing in clinical trials: An experience with two large cancer screening trials. PLoS medicine, 14(5), e1002304. doi:10.1371/journal.pmed.1002304

Versions

Version 3 (Current) : Updated 2021/09/24

Data Type	Download all or Query/Filter
CT Images (DICOM, 11.3 TB)	*Download* *Search* This link downloads the entire collection, which is quite large, as *legacy single frame images*. See the Detailed Description tab for options to download the collection in smaller chunks.

	(Download requires the NBIA Data Retriever)
Primary Tumor Tissue Slide Images (SVS, 775 GB)	*Download* *Search* ***Additional images are available***: See Detailed Description. (Download and apply the IBM-Aspera-Connect plugin to your browser to retrieve this faspex package)
Clinical data (ZIP, 25 MB)	*Download* (more info) Provided in SAS format in one compressed file (.zip); includes data and dictionaries. ***This is a subset of the full clinical data. If you need the full clinical data, please visit the Cancer Data Access System (CDAS) system.***

Data embargo of limited access is lifted September 2021, with the addition of downloadable pathology slide data and clinical data spreadsheet & dictionaries.

Version 2: Updated 2015/12/14

Data Type	Download all or Query/Filter
Images (DICOM, 11.3TB)	*Search*

Change: restoration of images that had become corrupted/missing during a storage transfer.

Version 1: Updated 2013/03/01

Data Type	Download all or Query/Filter
Images (DICOM, 11.3TB)	*Search*