

NLST Pathology

Redirection Notice

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

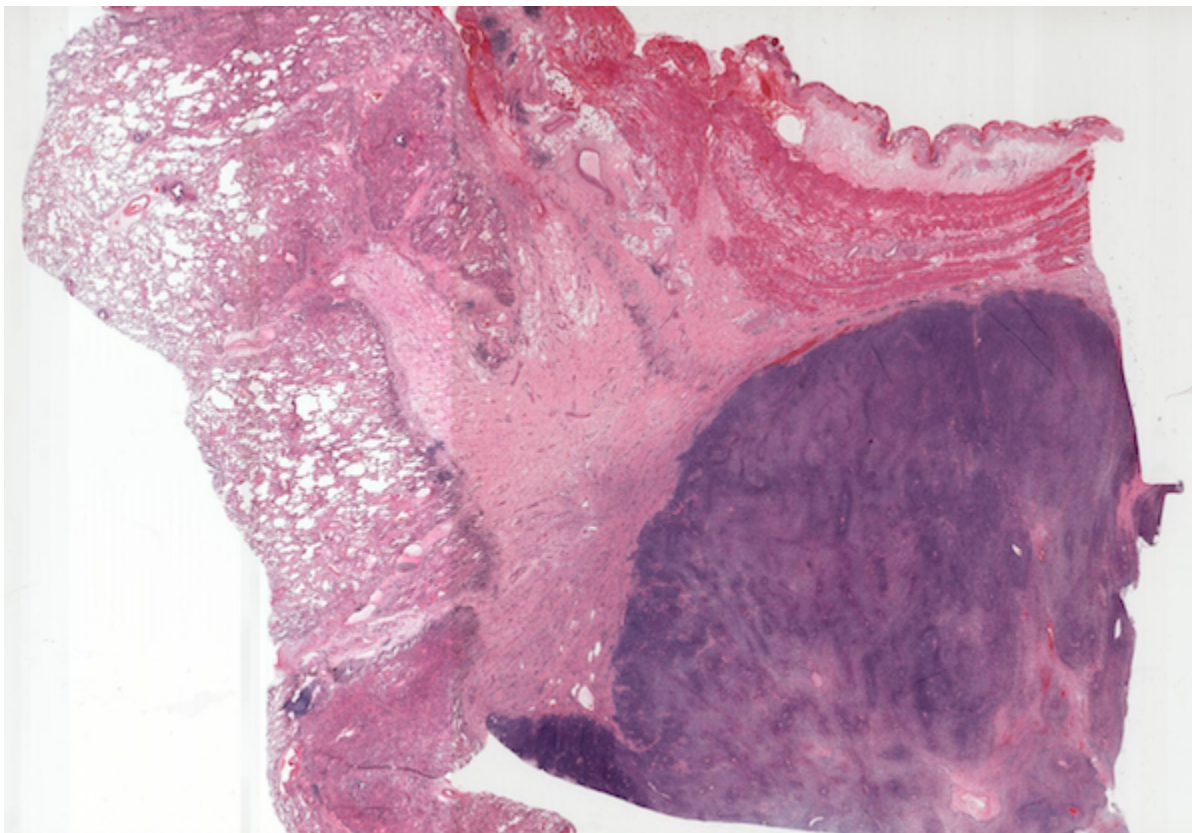
Summary

The Cancer Imaging Archive hosts pathology images from the [National Lung Screening Trial](#), in addition to the existing [radiology images](#) and a subset of the full clinical data. If you need the full clinical data, please visit the [Cancer Data Access System \(CDAS\)](#) system. The NLST was conducted by two separate networks of screening centers:

1) Lung Screening

Study Network (LSS): 10 centers and 34,612 participants; and 2) American College of Radiology Imaging Network (ACRIN): 23 centers and 18,840 participants. TCIA only hosts pathology data from LSS. The ACRIN pathology images and data are not available through TCIA.

As part of the pathology collection effort tissue, images, and data was obtained from 463 lung cancer patients (out of 1,284 total lung cancer patients in LSS). Note that nine of these 463 participants had two primary lung tumors; these are accessible through TCIA as an optional additional download for these nine participants.



About the Pathology Data




The NLST pathology specimen collection consists of biospecimens, images, and data for research. Near the conclusion of NLST activities, all available blocks of preserved lung tissue from NLST lung cancer patients were requested from pathology labs. Tissue cores were sampled from these donor blocks and placed into tissue microarrays (TMAs) and into Eppendorf tubes for analysis by the research community; the request process for the tissue specimens is under development as of October 2013. The images were obtained as an intermediate step in TMA construction. A thin section was cut from each donor tissue block, stained with hematoxylin and eosin (H & E), and imaged using an Aperio ScanScope.

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)	 <p>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.</p>  <p>(Download requires the NBIA Data Retriever)</p>	CC BY 4.0
Primary Tumor Tissue Slide Images (451 subjects, SVS, 775 GB)	<p>Download</p> <p>Additional images are available: See Detailed Description.</p> <p>(Download and apply the IBM-Aspera-Connect plugin to your browser)</p>	CC BY 4.0
Clinical data including data dictionaries (SAS, ZIP, 25 MB)	 <p>(more info)</p> <p>Provided in SAS format in one compressed file (.zip); includes data and dictionaries.</p> <p><i>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.</i></p>	CC BY 4.0
Additional histopathology slide images Table 1 for which the participants have no Baseline Questionnaire data (2 subjects, DOCX, 13 KB)	<p>Download</p>	CC BY 4.0
Additional histopathology slide images for which the participants have no Baseline Questionnaire data (2 subjects, 4 files, SVS)	<p>Download</p> <p>(Download and apply the IBM-Aspera-Connect plugin to your browser)</p>	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)	<p>Download</p>	CC BY 4.0
Additional histopathology 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)	<p>Download</p> <p>(Download and apply the IBM-Aspera-Connect plugin to your browser)</p>	CC BY 4.0

Detailed Description

Detailed Description

Collection Statistics	Radiology	Pathology
Modalities	CT (Legacy Single-Frame CT Images)	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 (GB)	11.3 TB	775 GB

More about NLST pathology slide data:

1. Primary Tumor slides (the standard package), 1225 files:
 - a. [link to faspex package](#)
 - b. [caMicroscope User Guide](#) for [caMicroscope](#), the pathology viewer that provides researchers with an HTML5 based web client that can be used to view a digitized pathology image at full resolution. Users can zoom in/out, and pan across the image. caMicroscope also allows one to create annotations, save them and retrieve annotations that were previously drawn.
2. Additional histopathology slide images [Table 1](#) for which the participants have no Baseline Questionnaire data (4 slides): [link to faspex package](#)
3. Additional histopathology slide images [Table 2](#) for participants with Second Primary Tumors as well as those included in the "standard" package (23 slides, 18.7 GB): [link to faspex package](#)

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.

More about the Clinical Data:

Please see <https://biometry.nci.nih.gov/cdas/learn/nlst/data-collected/> for more details about extensive NLST clinical data collection.

Clinical data here are 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.

Citations & Data Usage Policy

Citations & Data Usage Policy

Users of this data must abide by the [TCIA Data Usage Policy](#) and the [Creative Commons Attribution 4.0 International License](#) under which it has been published. 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: [10.1007/s10278-013-9622-7](https://doi.org/10.1007/s10278-013-9622-7)




Other Publications Using This Data

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](#).

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

Versions

Version 3 (Current) : Updated 2021/09/26

Data Type	Download all or Query/Filter
CT Images (DICOM, 11.3 TB)	<div>   </div> <p>(Download requires the NBIA Data Retriever)</p>
Tissue Slide Images (SVS, 775 GB)	<p>Download</p> <ul style="list-style-type: none"> Primary Tumor slides (the standard package), 1225 files: <link to faspex package> Additional histopathology slide images <link to Table 1> for which the participants have no Baseline Questionnaire data (4 slides): <link to additional faspex package> Additional histopathology slide images <link to Table 2> for participants with Second Primary Tumors (Primary for these are included in the "standard" package, description in Table 2 below) (23 slides): <link to additional faspex package> <p>(Download and apply the IBM-Aspera-Connect plugin to your browser)</p>
Clinical data (ZIP, 25 MB)	<div>  </div> <p>(more info)</p> <p>Provided in SAS format in one compressed file (.zip); includes data and dictionaries, other formats are available from IMS upon request.</p>

Data embargo of limited access is lifted September 2021, with the addition of (1) downloadable pathology slide data and (2) clinical data spreadsheet & dictionaries with no further restriction on number of participants per downloaded cohort.

Version 2: Updated 2015/12/14

Data Type	Download all or Query/Filter
Images (DICOM, 11.3TB)	<deprecated>

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)	<deprecated>