

NLST Participant: Data Dictionary

TABLE OF CONTENTS

Document Summary.....	2
NLST Participant: Data Dictionary	3
Section 1: Study	3
Section 2: Demographic.....	5
Section 3: Smoking	7
Section 4: Screening	8
Section 5: Follow-up / Procedures	10
Section 6: Lung Cancer	12
Section 7: Extended Follow-up Death.....	16
Section 8: Work History	17
Section 9: Disease History.....	20
Section 10: Personal Cancer History	22
Section 11: Family History of Lung Cancer	24
Section 12: Alcohol	25
Section 13: Abnormality Summary	27
Section 14: Progression	28
Section 15: SCT/Path Image Information	33
Section 16: Trial Last Contact	37
Section 17: Trial Death	39
Section 18: Coronary Artery Calcification.....	41
Section 19: Trial Endpoint Verification Process (EVP)	44

Document Summary

Property	Value
Document Title	NLST Participant: Data Dictionary
Date Created	02/17/2026
Sections	19
Entries	304
Document Filename	dictionary_prsn-20110203-021326.rtf

NLST Participant: Data Dictionary

Section 1: Study

Variable	Label	Description	Format Text
cen	Institution	A two character code, randomly assigned to each study center, in order to mask the identity of the centers.	Character
dataset_version	Date Stamp for Datasets		Char, 23
elig	Is participant eligible?	Was the participant eligible for the NLST trial? See the entry for "ineligible" for a list of eligibility criteria.	0="Ineligible Participant Randomized" 2="Eligible Participant"
ineligible	Reason for ineligibility	To be eligible, a participant must be 55-74 years of age on the date of randomization, a current smoker or a former smoker who quit smoking within the last 15 years, and have smoked at least 30 pack-years. Participants are ineligible if they do not meet those criteria, or if they meet any of the other exclusion criteria listed as reasons for ineligibility in the format for this variable.	.N="Not ineligible" 1="Age <55 or >74 yrs" 2="Non-smoker or quit > 15 years" 3="Insufficient pack years" 4="CT within 18 months of enrollment" 5="Participant in another cancer screening trial" 6="Participant in another cancer prevention trial" 7="Previous Lung Cancer" 8="Portion of lung removed" 9="Cancer within past 5 years" 10="Physical impairments to screening" 11="Metallic implants" 12="Home oxygen" 13="Unexplained weight loss or Hemoptysis" 14="Recent antibiotics use"
pid	Participant Identifier	A unique identifier given to each participant. For LSS participants, pid has a format of 1xx,xxx, while for ACRIN participants, pid has a format of 2xx,xxx.	Numeric
rndgroup	Study arm	The study arm to which the participant was randomized. Each participant was randomly assigned to either the spiral CT or X-ray arm. Because X-ray screening has been studied in other trials, the X-ray arm represents the control arm for determining the effectiveness of spiral CT screening.	1="Spiral CT" 2="X-ray"

Variable	Label	Description	Format Text
study	Which study is the participant enrolled on?	The component study to which the participant belonged. The National Lung Screening Trial (NLST) is composed of two component studies conducted by the Lung Screening Study group (LSS) and the American College of Radiology Imaging Network (ACRIN). ACRIN had some centers that collected biomarkers, while others did not.	1="LSS" 2="ACRIN biomarkers" 3="ACRIN no biomarkers"

Section 2: Demographic

Variable	Label	Description	Format Text
age	Age at randomization (in years; whole number)		Numeric
educat	Level of education completed		1="8th grade or less" 2="9th-11th grade" 3="High school graduate/GED" 4="Post high school training, excluding college" 5="Associate degree/ some college" 6="Bachelors Degree" 7="Graduate School" 8="Other" 95="Missing data form - form is not expected to ever be completed" 98="Missing - form was submitted and the answer was left blank" 99="Unknown/ decline to answer"
ethnic	Ethnicity	Is the participant of Hispanic/Latino heritage? Self-reported at entry.	1="Hispanic or Latino" 2="Neither Hispanic nor Latino" 7="Participant refused to answer" 95="Missing data form - form is not expected to ever be completed" 98="Missing - form was submitted and the answer was left blank" 99="Unknown/ decline to answer"
gender	Gender		1="Male" 2="Female"
height	Height in inches	Height of the participant in inches as self-reported at the start of the trial.	Numeric .M="Missing"
marital	Marital Status	Marital status of the participant at the start of the trial.	.M="Missing" 1="Never married" 2="Married or living as married" 3="Widowed" 4="Separated" 5="Divorced" 7="Participant refused to answer" 9="Not Ascertained"
race	Race		1="White" 2="Black or African-American" 3="Asian" 4="American Indian or Alaskan Native" 5="Native Hawaiian or Other Pacific Islander" 6="More than one race" 7="Participant refused to answer" 95="Missing data form - form is not expected to ever be completed" 96="Missing - no response" 98="Missing - form was submitted and the answer was left blank" 99="Unknown/ decline to answer"

Variable	Label	Description	Format Text
weight	Weight in pounds	Weight of the participant in pounds as self-reported at the start of the trial.	Numeric .M="Missing"
has_bq	Did the participant complete the baseline questionnaires?	Was the baseline medical history form administered for the participant?	0="No" 1="Yes"

Section 3: Smoking

Variable	Label	Description	Format Text
age_quit	Age at smoking cessation	Age at which a participant stopped smoking cigarettes.	Numeric .N="No age given"
cigar	Participant smokes/smoked cigars		.M="Missing" 0="No" 1="Yes"
cigsmok	Smoking status at T0	Cigarette smoking status (current vs former) at randomization. Former smokers must have quit within 15 years of eligibility determination to have been eligible for the trial.	0="Former" 1="Current"
pipe	Participant smokes/smoked a pipe		.M="Missing" 0="No" 1="Yes"
pkyr	Pack years	Pack years, calculated as: (Total Years Smoked x Cigarettes Per Day / 20). Allows for 2 significant digits.	Numeric
smokeage	Age at smoking onset	Age at which a participant started smoking cigarettes.	Numeric .M="Missing"
smokeday	Avg. num. of cigarettes per day		Numeric
smokelive	Participant lives/lived with smoker		.M="Missing" 0="No" 1="Yes"
smokework	Participant works/worked with exposure to smokers		.M="Missing" 0="No" 1="Yes"
smokeyr	Total years of smoking	Total number of years the participant smoked cigarettes. For LSS, this was calculated from the age of the participant at randomization (for current smokers) or the age the participant quit smoking (for former smokers), the age the participant began smoking, and the total number of years the participant did not smoke during that time. For ACRIN participants this is the total number of years the participant smoked as reported by the participant.	Numeric

Section 4: Screening

Variable	Label	Description	Format Text
scr_days0-2	Days since randomization at time of screening (for T0, T1, and T2 exams)	Days from randomization to the date of the screening exam. Contains information on either the spiral CT or the chest x-ray screening exam, depending on which study arm the participant was randomized to (see rndgroup).	Numeric .E="Screen date after lung cancer diagnosis" .N="No screen date on record" .W="Wrong Screen Administered"
scr_group	Screening Result Group		1="Screen-detected cancer" 2="Has nodule(s)" 3="Screened, but no nodules" 4="Never screened" 5="Other lung cancer"
scr_iso0-2	Result of isolation screen (for T0, T1, and T2 exams)	Isolation read of the image(s) from the screening exam before comparing with prior screening images or other available prior images. Contains information on either the spiral CT or the chest x-ray screening exam, depending on which study arm the participant was randomized to (see rndgroup).	1="Negative screen, no significant abnormalities" 2="Negative screen, minor abnormalities not suspicious for lung cancer" 3="Negative screen, significant abnormalities not suspicious for lung cancer" 4="Positive, Change Unspecified, nodule(s) >= 4 mm or enlarging nodule(s), mass(es), other non-specific abnormalities suspicious for lung cancer" 10="Inadequate Image" 11="Not Compliant - Left Study" 13="Not Expected - Cancer before screening window" 14="Not Expected - Death before screening window" 15="Not Compliant - Refused a screen" 17="Not Compliant - Wrong Screen" 23="Not Expected - Cancer in screening window" 24="Not Expected - Death in screening window" 95="Not Compliant - Erroneous Report of Lung Cancer Before Screen (LSS Only)" 97="Not Compliant - Form Not Submitted, Window Closed"
scr_lat0-2	Did the participant receive a lateral chest x-ray (in study years T0, T1, and T2)?	Did the participant receive a lateral chest x-ray instead of the proper screening exam? Participants in the x-ray arm were supposed to receive posteroanterior chest x-rays, but a small number of participants were incorrectly given lateral chest x-rays.	0="Did not receive lateral view chest x-ray (includes CT, X-Ray without lateral, no screen, etc) " 1="Received lateral view chest x-ray"

Variable	Label	Description	Format Text
scr_res0-2	Results of screening (for T0, T1, and T2 exams)	Official result of the screening exam, after comparing the current image(s) with prior screening images or other available prior images. Participants with positive screens (scr_resX = 4, 5, or 6) were advised to receive a diagnostic evaluation for lung cancer with their personal physician. Contains information on either the spiral CT or the chest x-ray screening exam, depending on which study arm the participant was randomized to (see rndgroup).	1="Negative screen, no significant abnormalities" 2="Negative screen, minor abnormalities not suspicious for lung cancer" 3="Negative screen, significant abnormalities not suspicious for lung cancer" 4="Positive, Change Unspecified, nodule(s) >= 4 mm or enlarging nodule(s), mass(es), other non-specific abnormalities suspicious for lung cancer" 5="Positive, No Significant Change, stable abnormalities potentially related to lung cancer, no significant change since prior screening exam." 6="Positive, other" 10="Inadequate Image" 11="Not Compliant - Left Study" 13="Not Expected - Cancer before screening window" 14="Not Expected - Death before screening window" 15="Not Compliant - Refused a screen" 17="Not Compliant - Wrong Screen" 23="Not Expected - Cancer in screening window" 24="Not Expected - Death in screening window" 95="Not Compliant - Erroneous Report of Lung Cancer Before Screen (LSS Only)" 97="Not Compliant - Form Not Submitted, Window Closed"
last_screen_studyyr	Study year of the last screening exam	The study year of the last screen the participant received.	.N="No Screen" 0="Study year 0" 1="Study year 1" 2="Study year 2"
num_screens	Count of screening rounds completed	Number of screening rounds with a valid, adequate result. This only includes screens with a positive or negative result (scr_res of 1-6). Inadequate screens and screens that are not considered valid (e.g., because they occurred after lung cancer diagnosis) are not included in this count.	Numeric

Section 5: Follow-up / Procedures

Variable	Label	Description	Format Text
biop0-2	Had a biopsy related to positive screen?	Was a biopsy performed as followup to a positive screen at T0, T1, or T2?	0="No" 1="Yes"
bioplc	Had a biopsy related to lung cancer?	Was a biopsy performed in relation to the diagnosis of lung cancer?	0="No" 1="Yes"
invas0-2	Had an invasive procedure related to positive screen?	Was an invasive procedure done as followup to a positive screen at T0, T1, or T2? Invasive procedures include thoracotomy, thoracoscopy, mediastinotomy, mediastinoscopy, and biopsies.	0="No" 1="Yes"
invaslc	Had an invasive procedure related to lung cancer?	Was an invasive procedure done in relation to the diagnosis of lung cancer? Invasive procedures include thoracotomy, thoracoscopy, mediastinotomy, mediastinoscopy, and biopsies.	0="No" 1="Yes"
medcomp0-2	Had any complications related to positive screen?	Was there a medical complication in relation to the followup of a positive screen at T0, T1, or T2?	0="No" 1="Yes"
medcomplc	Had any complications related to lung cancer?	Was there a medical complication in relation to the diagnosis of lung cancer?	0="No" 1="Yes"
mra_stat0-2	Positive screen follow-up status by study year (from Medical Records Abstraction)	What level of followup was done following a positive screen at T0, T1, or T2?	0="No Expectation" 1="MRA Complete with Diagnostic Procedures" 2="MRA Complete with clinical evaluation/radiograph-comparison with historical images only" 3="MRA Complete with no procedures" 4="MRA Complete, procedures indeterminate" 5="Conflicting Data"
no_proc_reas0-2	Reason for no procedures when follow-up was expected	What is the reason no followup procedures were done following a positive screen at T0, T1, or T2?	0="Either had follow-up to positive screen or no follow-up was expected" 1="Provider/Radiologist did not recommend follow-up" 2="Participant declined to undergo follow-up" 3="No follow-up for other reasons" 4="No follow-up and reason unknown"
proc0-2	Had a procedure related to positive screen?	Was any diagnostic procedure done as followup to a positive screen at T0, T1, or T2?	0="No" 1="Yes"

Variable	Label	Description	Format Text
proc1c	Had any procedure related to lung cancer?	Was any diagnostic procedure done in relation to the diagnosis of lung cancer?	0="No" 1="Yes"

Section 6: Lung Cancer

Variable	Label	Description	Format Text
can_scr	Result of screen associated with the first confirmed lung cancer diagnosis	Indicates whether the cancer followed a positive, negative, or missed screen, or whether it occurred after the screening years.	0="No Cancer" 1="Positive Screen" 2="Negative Screen" 3="Missed Screen" 4="Post Screening"
canc_free_days	Days from randomization to date when participant was last known to be free from lung cancer	Days until the date the participant was last known to be free of lung cancer. This date comes from a participant study update form, except for ACRIN participants with cancer, where the cancer diagnosis date is used. Do not use this variable to calculate follow-up time for lung cancer incidence rates, Cox regressions, etc. Instead, use fup_days for non-cases and candx_days for cases.	Numeric
canc_rpt_link	Is the diagnosis of lung cancer associated with a positive screen?	Is the diagnosis of lung cancer associated with a positive screen? A positive screen and a cancer diagnosis are associated based on a linking algorithm using diagnostic procedures.	0="No" 1="Yes"
canc_rpt_source	Source of lung cancer report	Source of the information that prompted the medical record abstracting to confirm the lung cancer. For LSS participants, the level 5 indicates that cancer was confirmed during the medical record abstraction process after a positive screen. For ACRIN participants, cancers found from positive screens are placed in the level 0, and the level 5 indicates that it was determined the participant had cancer during medical record abstraction for other purposes besides confirming lung cancer.	0="No Report" 1="Participant Self Report (ASU or F1/F2)" 2="Death Certificate" 3="Other Report (CC/CNF)" 4="Endpoint Verification Process (EVP)" 5="Medical Records Abstraction" 7="National Death Index Search (NDI)"

Variable	Label	Description	Format Text
cancyr	Study year associated with first confirmed lung cancer	Study year associated with the first confirmed lung cancer. A cancer associated with a positive screen is assigned the study year of that screen.	.N="Not Applicable" 0="T0" 1="T1" 2="T2" 3="T3" 4="T4" 5="T5" 6="T6" 7="T7"
candx_days	Days from randomization to first diagnosis of lung cancer		Numeric .N="No diagnosis date on record"
confic	Status of lung cancer report	The result of the medical record abstracting that attempted to confirm the report of lung cancer.	0="No Report" 1="Follow-up collected - Confirmed Lung Cancer" 2="Follow-up Collected - Confirmed Not Lung Cancer" 3="Medical Records cannot be obtained" 4="Pending"
de_grade	Lung cancer grade		.N="Not Applicable" 1="Grade Cannot Be Assessed (GX) " 2="Well Differentiated (G1)" 3="Moderately Differentiated (G2)" 4="Poorly Differentiated (G3)" 5="Undifferentiated (G4)" 6="Unspecified in Pathology Report" 8="Unknown" 9="Missing"
de_stag	Lung cancer Stage	Lung cancer stage (AJCC 6), combining clinical and pathologic staging information.	.M="Missing" .N="Not Applicable" 110="Stage IA" 120="Stage IB" 210="Stage IIA" 220="Stage IIB" 310="Stage IIIA" 320="Stage IIIB" 400="Stage IV" 888="TNM not available" 900="Occult Carcinoma" 994="Carcinoid, cannot be assessed" 999="Unknown, cannot be assessed"
de_stag_7thed	AJCC 7th edition stage	Stage of first primary lung cancer, based on AJCC 7th edition stage	.M="Missing" .N="Not Applicable" 110="Stage IA" 120="Stage IB" 210="Stage IIA" 220="Stage IIB" 310="Stage IIIA" 320="Stage IIIB" 400="Stage IV" 900="Occult Carcinoma" 999="Unknown, cannot be assessed"
de_type	Lung cancer type from ICD-O-3 morphology	Lung cancer type from ICD-O-3 morphology. For LSS participants, this is recorded separately from the complete ICD code, and represents the best information available on the type of the cancer.	Numeric .M="Missing" .N="Not Applicable"

Variable	Label	Description	Format Text
lesionsize	Pathology lesion size of tumor in millimeters		Numeric .M="Missing" .N="Not Applicable"
loccar	Cancer in Carina	Was the primary tumor located in the carina? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
loclhil	Cancer in Left Hilum	Was the primary tumor located in the left hilum? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
loclin	Cancer in Lingula	Was the primary tumor located in the lingula? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locllow	Cancer in Left lower lobe	Was the primary tumor located in the left lower lobe? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locmsb	Cancer in Left main stem bronchus	Was the primary tumor located in the left main stem bronchus? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
loclup	Cancer in Left upper lobe	Was the primary tumor located in the left upper lobe? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locmed	Cancer in Mediastinum	Was the primary tumor located in the mediastinum? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locoth	Cancer in Other Location	Was the primary tumor located in an other specified location? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locrhil	Cancer in Right Hilum	Was the primary tumor located in the right hilum? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locrlow	Cancer in Right lower lobe	Was the primary tumor located in the right lower lobe? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locrmid	Cancer in Right middle lobe	Was the primary tumor located in the right middle lobe? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"

Variable	Label	Description	Format Text
locrmsb	Cancer in Right main stem bronchus	Was the primary tumor located in the right main stem bronchus? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locrup	Cancer in Right upper lobe	Was the primary tumor located in the right upper lobe? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
locunk	Cancer in Unknown location	Was the primary tumor located in an unknown location? A tumor may be located in more than one location.	.N="Not Applicable" 0="No" 1="Yes"
treatlc	Status of treatment data	Did the participant receive treatment for lung cancer?	.N="Not Applicable" 1="Confirmed treatment" 2="Confirmed no treatment" 3="Treatment data incomplete"
histology_cat	Categorization of lung cancer type from ICD-O-3 morphology	Categorization of ICD-O-3 morphology. For LSS participants, this is separate from the morphology recorded with the complete ICD code, and represents the best information known about the participant's cancer.	.M="Missing" .N="Not Applicable" 1="Small cell carcinoma" 2="Squamous cell carcinoma" 3="Adenocarcinoma" 4="Bronchiolo-alveolar carcinoma" 5="Large cell carcinoma" 6="Adenosquamous carcinoma" 7="Pleomorphic/sarcomatoid" 8="Carcinoid tumor" 9="Unclassified carcinoma"
lung_cancer	Confirmed Lung Cancer?	Does the participant have a confirmed lung cancer diagnosis?	0="No" 1="Yes"

Section 7: Extended Follow-up Death

Variable	Label	Description	Format Text
death_days_2015	Days since randomization at death, including deaths from 2015 NDI		Numeric .N="No date of death on record"
deathicd_2015	Cause of death, including deaths from 2015 NDI		Character " "="Missing"
finaldeathlc_2015	Is lung cancer the cause of death, including deaths from 2015 NDI	This is the authoritative variable indicating whether lung cancer was the cause of death. It uses the information from the Endpoint Verification Process (EVP) if available and uses the death certificate if EVP is not available. This variable does not reflect updated death data for all centers: one study center with fewer than 1200 participants did not submit participants for the updated NDI search.	.M="Death Reported, cause of death unknown" .N="No death reported" 0="Death not due to lung cancer" 1="Death due to lung cancer or work-up of suspected lung cancer"
is_dead_2015	Is the Participant Dead? (Including deaths from 2015 NDI)	Is the participant considered to be dead? This variable does not reflect updated death data for all centers: one study center with fewer than 1200 participants did not submit participants for the updated NDI search.	0="No" 1="Yes"
mortality_exitdays_2015	Days from Randomization to Mortality Exit (Including Data from 2015 NDI)		Numeric
mortality_exitstat_2015	Exit Status for Mortality (Including Data from 2015 NDI)	Exit status for mortality. This variable does not reflect updated death data for all centers: one study center with fewer than 1200 participants did not submit participants for the updated NDI search.	1="Death" 2="Date known alive from NDI" 3="Other date known alive"

Section 8: Work History

Variable	Label	Description	Format Text
resasbe	Asbestos work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
resbaki	Baking work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes"
resbutc	Butchering/meat packing work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
reschem	Chemicals/plastics manufacturing work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
rescoal	Coal mining work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
rescott	Cotton/jute processing work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
resfarm	Farming work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
resfire	Fire fighting work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
resflou	Flour/feed or grain milling work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
resfoun	Foundry/steel milling work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"

Variable	Label	Description	Format Text
reshard	Hard rock mining work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
respain	Painting work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
ressand	Sandblasting work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
resweld	Welding work: Wore mask or other equipment to protect lungs?		.M="Missing" .N="Not Applicable" 0="No" 1="Yes" 2="Sometimes"
wrkasbe	Asbestos: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkbaki	Baking: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkbutc	Butchering/meat packing: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkchem	Chemicals/plastics manufacturing: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkcoal	Coal mining: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkcott	Cotton/jute processing: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkfarm	Farming: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkfire	Fire fighting: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkflou	Flour/feed or grain milling: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkfoun	Foundry/steel milling: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkhard	Hard rock mining : Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"

Variable	Label	Description	Format Text
wrkpain	Painting: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrksand	Sandblasting: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
wrkweld	Welding: Ever worked for 1 or more years?		.M="Missing" 0="No" 1="Yes"
yrsasbe	Asbestos work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrsbaki	Baking work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrsbutc	Butchering/meat packing work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrschem	Chemicals/plastics manufacturing work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrscoal	Coal mining work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrscott	Cotton/jute processing work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrsfarm	Farming work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrsfire	Fire fighting work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrsflou	Flour/feed or grain milling work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrsfoun	Foundry/steel milling work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrshard	Hard rock mining work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrspain	Painting work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrs sand	Sandblasting work: Number of years?		Numeric .M="Missing" .N="Not Applicable"
yrs weld	Welding work: Number of years?		Numeric .M="Missing" .N="Not Applicable"

Section 9: Disease History

Variable	Label	Description	Format Text
ageadas	Asthma (adult): Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
ageasbe	Asbestosis: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agebron	Bronchiectasis: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agechas	Asthma (childhood): Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agechro	Chronic bronchitis: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agecopd	COPD: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agediab	Diabetes: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
ageemph	Emphysema: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agefibr	Fibrosis of the lung: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agehear	Heart disease or heart attack: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agehype	Hypertension: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agepneu	Pneumonia: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agesarc	Sarcoidosis: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agesili	Silicosis: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agestro	Stroke: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"

Variable	Label	Description	Format Text
agetube	Tuberculosis: Age at diagnosis prior to trial?		Numeric .M="Missing" .N="Not Applicable"
diagadas	Asthma (adult): Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagasbe	Asbestosis: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagbron	Bronchiectasis: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagchas	Asthma (childhood): Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagchro	Chronic bronchitis: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagcopd	COPD: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagdiab	Diabetes: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagemph	Emphysema: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagfibr	Fibrosis of the lung: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diaghear	Heart disease or heart attack: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diaghype	Hypertension: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagpneu	Pneumonia: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagsarc	Sarcoidosis: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagsili	Silicosis: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagstro	Stroke: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
diagtube	Tuberculosis: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"

Section 10: Personal Cancer History

Variable	Label	Description	Format Text
ageblad	Bladder Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agebrea	Breast Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agecerv	Cervical Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agecolo	Colorectal Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
ageesop	Esophageal Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agekidn	Kidney Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agelary	Larynx Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agelung	Lung Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agenasa	Nasal Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
ageoral	Oral Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agepanc	Pancreatic Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agephar	Pharynx Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agestom	Stomach Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agethyr	Thyroid Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"
agetran	Transitional Cell Cancer: Age at diagnosis prior to trial		Numeric .M="Missing" .N="Not Applicable"

Variable	Label	Description	Format Text
cancblad	Bladder Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
cancbrea	Breast Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc cerv	Cervical Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canccolo	Colorectal Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
cancesop	Esophageal Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc kidn	Kidney Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc lary	Larynx Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc lung	Lung Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc nasa	Nasal Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
cancoral	Oral Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc panc	Pancreatic Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
cancphar	Pharynx Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc stom	Stomach Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc thy	Thyroid Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"
canc tran	Transitional Cell Cancer: Ever diagnosed prior to trial?		.M="Missing" 0="No" 1="Yes"

Section 11: Family History of Lung Cancer

Variable	Label	Description	Format Text
fambrother	Did brother(s) (including half-brothers) have lung cancer?		.M="Missing" 0="No" 1="Yes"
famchild	Did child (biological) have lung cancer?		.M="Missing" 0="No" 1="Yes"
famfather	Did father have lung cancer?		.M="Missing" 0="No" 1="Yes"
fammother	Did mother have lung cancer?		.M="Missing" 0="No" 1="Yes"
famsister	Did sister(s) (including half-sisters) have lung cancer?		.M="Missing" 0="No" 1="Yes"

Section 12: Alcohol

Variable	Label	Description	Format Text
acrin_alc_curr	Do you presently drink alcoholic beverages?	For ACRIN participants, was the participant a current drinker at time of entry.	.M="Missing" .N="Not Applicable" .R="Participant Refused to Answer" 1="No" 2="Yes"
acrin_alc_ever	Have you ever consumed alcoholic beverages?	For ACRIN participants, did the participant ever drink alcohol.	.M="Missing" .N="Not Applicable" .R="Participant Refused to Answer" 1="No" 2="Yes"
acrin_drink24hr	During the past 24 hours, how many drinks have you had?	For ACRIN participants who are current drinkers, how many drinks did the participant have in the last 24 hours at time of entry.	Numeric .M="Missing" .N="Not Applicable" .R="Participant Refused to Answer"
acrin_drinknum_curr	What is the usual number of drinks you have per week?	For ACRIN participants who are current drinkers, the average number of drinks the participant had per week at time of entry.	Numeric .M="Missing" .N="Not Applicable" .R="Participant Refused to Answer"
acrin_drinknum_form	What was the usual number of drinks you had per week?	For ACRIN participants who are former drinkers, the average number of drinks the participant had per week before they stopped drinking.	Numeric .M="Missing" .N="Not Applicable" .R="Participant Refused to Answer"
acrin_drinkyrs_curr	For how many years have you been drinking alcoholic beverages?	For ACRIN participants who are current drinkers, how many years they have been drinking alcohol.	Numeric .M="Missing" .N="Not Applicable" .R="Participant Refused to Answer"
acrin_drinkyrs_form	For how many years did you drink alcoholic beverages?	For ACRIN participants who are former drinkers, how many years they drank alcohol.	Numeric .M="Missing" .N="Not Applicable" .R="Participant Refused to Answer"
acrin_lastdrink	How long has it been since you last had an alcoholic drink?	For ACRIN participants who are former drinkers, how long it has been since they last drunk alcohol.	.M="Missing" .N="Not Applicable" .R="Participant Refused to Answer" 1="Less than 1 year" 2="1 to 2 years" 3="More than 2 years"
lss_alcohol_freq	How often do you have a drink containing alcohol?		.M="Missing" .N="Not Applicable" 1="Never" 2="Monthly or less often" 3="Two to four times a month" 4="Two to three times a week" 5="Four or more times a week"

Variable	Label	Description	Format Text
lss_alcohol_num	Number of alcoholic drinks on typical day when drinking		.M="Missing" .N="Not Applicable" 1="1" 2="2-3" 3="4" 4="5-7" 5="8 or more"

Section 13: Abnormality Summary

Variable	Label	Description	Format Text
anyscr_has_nodule	Was a non-calcified nodule/mass >=4mm found on any screen?		.N="Not Applicable" 0="No" 1="Yes"

Section 14: Progression

Variable	Label	Description	Format Text
last_progfree_days	Days from randomization to last date known free of lung cancer progression	The number of days from randomization until a participant was last known to be free of lung cancer progression, for participants with lung cancer. For participants who had progression, this is the number of days until the date of progression. For participants who had no progression, this is the days until the end of the time period covered by their progression forms. For participants unknown for progression, this is the days until the start of the time period in which they were unknown for progression. This variable is set to missing for participants who never had a progression form completed.	Numeric .F="No Form" .N="Not Applicable"
prog_days_1st	Days from randomization to first lung cancer progression		Numeric .F="No Form" .N="Not Applicable"
prog_days_2nd	Days from randomization to second lung cancer progression		Numeric .F="No Form" .N="Not Applicable"
prog_days_3rd	Days from randomization to third lung cancer progression		Numeric .F="No Form" .N="Not Applicable"
prog_days_4th	Days from randomization to fourth lung cancer progression		Numeric .F="No Form" .N="Not Applicable"
prog_days_5th	Days from randomization to fifth lung cancer progression		Numeric .F="No Form" .N="Not Applicable"
progressed_ever	Ever had lung cancer progression		.F="No Form" .N="Not Applicable" 0="No Progression" 1="Had Progression" 9="Unknown if Had Progression"
progression_num	Number of times lung cancer progressed		Numeric .F="No Form" .N="Not Applicable"

Variable	Label	Description	Format Text
progsite_adrenal_1st	First lung cancer progression included progression to adrenal glands		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_adrenal_days	Days from randomization to first lung cancer progression to adrenal		Numeric .F="No Form" .N="Not Applicable"
progsite_adrenal_ever	Lung cancer ever progressed to adrenal glands		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_adrenal_num	Number of times lung cancer progressed to adrenal		Numeric .F="No Form" .N="Not Applicable"
progsite_bone_1st	First lung cancer progression included progression to bone		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_bone_days	Days from randomization to first lung cancer progression to bone		Numeric .F="No Form" .N="Not Applicable"
progsite_bone_ever	Lung cancer ever progressed to bone		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_bone_num	Number of times lung cancer progressed to bone		Numeric .F="No Form" .N="Not Applicable"
progsite_brain_1st	First lung cancer progression included progression to brain		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_brain_days	Days from randomization to first lung cancer progression to brain		Numeric .F="No Form" .N="Not Applicable"
progsite_brain_ever	Lung cancer ever progressed to brain		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_brain_num	Number of times lung cancer progressed to brain		Numeric .F="No Form" .N="Not Applicable"
progsite_liver_1st	First lung cancer progression included progression to liver		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_liver_days	Days from randomization to first lung cancer progression to liver		Numeric .F="No Form" .N="Not Applicable"
progsite_liver_ever	Lung cancer ever progressed to liver		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"

Variable	Label	Description	Format Text
progsite_liver_num	Number of times lung cancer progressed to liver		Numeric .F="No Form" .N="Not Applicable"
progsite_lymph_n1_1st	First lung cancer progression included progression to N1 lymph nodes		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_lymph_n1_days	Days from randomization to first lung cancer progression to N1 lymph nodes		Numeric .F="No Form" .N="Not Applicable"
progsite_lymph_n1_ever	Lung cancer ever progressed to N1 lymph nodes		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_lymph_n1_num	Number of times lung cancer progressed to N1 lymph nodes		Numeric .F="No Form" .N="Not Applicable"
progsite_lymph_n2_1st	First lung cancer progression included progression to N2 lymph nodes		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_lymph_n2_days	Days from randomization to first lung cancer progression to N2 lymph nodes		Numeric .F="No Form" .N="Not Applicable"
progsite_lymph_n2_ever	Lung cancer ever progressed to N2 lymph nodes		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_lymph_n2_num	Number of times lung cancer progressed to N2 lymph nodes		Numeric .F="No Form" .N="Not Applicable"
progsite_lymph_n3_1st	First lung cancer progression included progression to N3 lymph nodes		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_lymph_n3_days	Days from randomization to first lung cancer progression to N3 lymph nodes		Numeric .F="No Form" .N="Not Applicable"
progsite_lymph_n3_ever	Lung cancer ever progressed to N3 lymph nodes		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_lymph_n3_num	Number of times lung cancer progressed to N3 lymph nodes		Numeric .F="No Form" .N="Not Applicable"
progsite_mediastinum_1st	First lung cancer progression included progression to mediastinum		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"

Variable	Label	Description	Format Text
progsite_mediastinum_days	Days from randomization to first lung cancer progression to mediastinum		Numeric .F="No Form" .N="Not Applicable"
progsite_mediastinum_ever	Lung cancer ever progressed to mediastinum		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_mediastinum_num	Number of times lung cancer progressed to mediastinum		Numeric .F="No Form" .N="Not Applicable"
progsite_orig_lung_1st	First lung cancer progression included progression at original lung site		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_orig_lung_days	Days from randomization to first lung cancer progression at original lung site		Numeric .F="No Form" .N="Not Applicable"
progsite_orig_lung_ever	Lung cancer ever progressed at original lung site		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_orig_lung_num	Number of times lung cancer progressed at original lung site		Numeric .F="No Form" .N="Not Applicable"
progsite_other_1st	First lung cancer progression included progression to other site		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_other_days	Days from randomization to first lung cancer progression to other site		Numeric .F="No Form" .N="Not Applicable"
progsite_other_ever	Lung cancer ever progressed to other site		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_other_lung_1st	First lung cancer progression included progression to other lung site		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_other_lung_days	Days from randomization to first lung cancer progression to other lung site		Numeric .F="No Form" .N="Not Applicable"
progsite_other_lung_ever	Lung cancer ever progressed to other lung site		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_other_lung_num	Number of times lung cancer progressed to other lung site		Numeric .F="No Form" .N="Not Applicable"

Variable	Label	Description	Format Text
progsite_other_num	Number of times lung cancer progressed to other site		Numeric .F="No Form" .N="Not Applicable"
progsite_pleura_1st	First lung cancer progression included progression to pleura		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_pleura_days	Days from randomization to first lung cancer progression to pleura		Numeric .F="No Form" .N="Not Applicable"
progsite_pleura_ever	Lung cancer ever progressed to pleura		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_pleura_num	Number of times lung cancer progressed to pleura		Numeric .F="No Form" .N="Not Applicable"
progsite_skin_1st	First lung cancer progression included progression to skin		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_skin_days	Days from randomization to first lung cancer progression to skin		Numeric .F="No Form" .N="Not Applicable"
progsite_skin_ever	Lung cancer ever progressed to skin		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_skin_num	Number of times lung cancer progressed to skin		Numeric .F="No Form" .N="Not Applicable"
progsite_unk_1st	First lung cancer progression included progression to unknown site		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_unk_days	Days from randomization to first lung cancer progression to unknown site		Numeric .F="No Form" .N="Not Applicable"
progsite_unk_ever	Lung cancer ever progressed to unknown site		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
progsite_unk_num	Number of times lung cancer progressed to unknown site		Numeric .F="No Form" .N="Not Applicable"

Section 15: SCT/Path Image Information

Variable	Label	Description	Format Text
all_sct_batch	Full SCT Image Batch	Batch number associated with SCT image series. The full NLST CT image collection is divided into two sets which are, in turn, divided into 10 batches, each.	Numeric .I="No image series, sct_image_years = 0" .N="Not Applicable"
all_sct_set	Full SCT Image Set	Set number associated with the SCT image series in the full NLST SCT collection. The NLST SCT Collection was divided into two equally-distributed sets that were intended to be used as training and testing sets.	.I="No image series, sct_image_years = 0" .N="Not Applicable" 1="Set 1" 2="Set 2"
has_axial0-2	Has axial image series in Tx?		.I="No image series, sct_image_years = 0" .N="Not applicable" .S="No image series for study year, sct_image_years > 0" 0="No" 1="Yes"
has_axial_any	Has axial image series in any study year?		.I="No image series, sct_image_years = 0" .N="Not applicable" 1="Yes"
num_axial_series0-2	Number of axial series in Tx		Numeric .I="No image series, sct_image_years = 0" .N="Not applicable" .S="No image series for study year, sct_image_years > 0"
num_image_series0-2	Number of image series in Tx		Numeric .I="No image series, sct_image_years = 0" .N="Not applicable" .S="No image series for study year, sct_image_years > 0"
num_localizer_series0-2	Number of localizer series in Tx		Numeric .I="No image series, sct_image_years = 0" .N="Not applicable" .S="No image series for study year, sct_image_years > 0"

Variable	Label	Description	Format Text
scr_group_2021	CT Image Population (2021)	This variable defines a value that is used to categorize participants for CT image selections that are focused on positive screens and lung cancers. The variable is based on CT image data received in 2021. Categorizations are based on participants' lung cancer status, nodule status, and axial image status. For screen-detected lung cancers, participants are required to have an axial image in the year of diagnosis. For screen-detected cancers, positive screens, and nodule positive participants, axial images are required in the years in which the outcome was noted.	.N="Not Applicable" 1="Screen Detected Lung Cancer & Image same Study Year" 2="Lung Cancer & Image" 3="Nodule positive OR Positive Screen & Image from same Study Year" 4="3 Negative Screens & 3 images" 5="Other non-eligible participants" 7="Participants with NO Images"
sct_axial_years	Number of study years with axial image series		Numeric .I="No image series, sct_image_years = 0" .N="Not applicable"
selected_lc_pop	Participant's SCT image series are included in the Lung Cancer Selection?	Is the image series included in the Lung Cancer Selection? The Lung Cancer Selection provides investigators with a set of images and data that are designed to provide the participants with lung cancers (screen-detected and otherwise) and a cross-section of control participants at a size that is manageable for download. The Lung Cancer Selection is divided into two sets which are, in turn, divided into 7 batches, each.	.I="No image series, sct_image_years = 0" .N="Not Applicable" 0="No" 1="Yes"

Variable	Label	Description	Format Text
selected_lc_pop_batch	Lung Cancer Selection Batch	Batch number associated with SCT image series in the Lung Cancer Selection. The Lung Cancer Selection provides investigators with a set of images and data that are designed to provide the participants with lung cancers (screen-detected and otherwise) and a cross-section of control participants at a size that is manageable for download. The Lung Cancer Selection is divided into two sets which are, in turn, divided into 7 batches, each.	.I="No image series, sct_image_years = 0" .N="Not Applicable" 1="A. Screen-detected lung cancers" 2="B. Other Lung Cancers" 3="C. Positive Screens" 4="D. Positive Screens" 5="E. Positive Screens" 6="F. Positive Screens" 7="G. Negative Screens"
selected_lc_pop_set	Lung Cancer Selection Set	Set number associated with the SCT image series in the Lung Cancer Selection The Lung Cancer Selection was divided into two equally-distributed sets that were intended to be used as training and testing sets. The Lung Cancer Selection provides investigators with a set of images and data that are designed to provide the participants with lung cancers (screen-detected and otherwise) and a cross-section of control participants at a size that is manageable for download. The Lung Cancer Selection is divided into two sets which are, in turn, divided into 7 batches, each.	.I="No image series, sct_image_years = 0" .N="Not Applicable" 1="Set 1" 2="Set 2"
selected_path	Selected for Standard Pathology Selection	Was the participant included in the Standard Pathology Selection? Participants are included in the Standard Pathology Selection if they have a path H&E image for their first lung primary and completed the baseline questionnaire.	0="No" 1="Yes"

Variable	Label	Description	Format Text
selected_path_batch	Standard Pathology Selection Image Batch	Batch number associated with path H&E images selected for inclusion in the Standard Pathology Selection.	.N="Not Applicable" 1="Adenocarcinoma" 2="Adenocarcinoma" 3="Adenocarcinoma" 4="Adenocarcinoma" 5="Squamous cell carcinoma" 6="Squamous cell carcinoma" 7="Squamous cell carcinoma" 8="Bronchio-alveolar carcinoma" 9="Other lung cancers"
selected_path_image_count	Number Of Path H&E Images	Number of path H&E images for participants in the Standard Pathology Selection. Only images for the first primary lung cancer are included.	Numeric

Section 16: Trial Last Contact

Variable	Label	Description	Format Text
contactstatus	Status of participant at 12/31/2009	Participation level at the end of the study (12/31/2009). Note that a small subset (< 1%) of participants who are "Active" have a last follow-up date before 12/31/2009. Some of them were considerably earlier than 12/31/2009. While these participants are classified as "Active," their vital status is only known through their last follow-up date.	1="Active" 2="Deceased" 3="Randomized but Never Participated" 4="Withdrawn or Lost Contact from Study"
fup_days	Days from randomization to date last known alive	Days from randomization until the date the participant was last known to be alive. For LSS, this is calculated as follows: 1. Death date from death certificate 2. Withdrawal date 3. Latest ASU date 4. Latest screen/randomization date For ACRIN, this must be combined from various sources: 1. Death date from death certificate (D1) 2. Death date reported on FC 3. Withdrawal date 4. Latest interval date on F1/FC 5. Latest screen/randomization date This is the appropriate variable to use for follow-up time for analyses of mortality and survival. For lung cancer incidence, use fup_days for non-cases and candx_days for cases.	Numeric

Variable	Label	Description	Format Text
wdlost	Reason why case is withdrawn or lost contact	Reason why non-deceased participants stopped actively participating in NLST.	0="Not Withdrawn or Lost Contact" 1="Participant Withdrew Consent" 2="Participant Refused Further Participation for Non-Medical Reasons" 3="Participant Refused Further Participation Due to Physical Illness/Cognitive Impairment" 4="Lost Contact/Cannot Locate/No Active Contact with Participant" 5="Administrative Withdrawal"

Section 17: Trial Death

Variable	Label	Description	Format Text
dcfdeathlc	Death certificate lung cancer death	This variable indicates whether the underlying cause of death from the death certificate is lung cancer (first 3 digits of DCFICD are "C34").	.M="Death Reported, DC cause of death unknown" .N="No death reported" 0="Death not due to lung cancer" 1="Death due to lung cancer"
dcfcd	ICD-10 code for the underlying cause of death	This variable provides a specific cause of death for all deceased participants with a death certificate. It is derived from multiple fields on the death certificate, according to rules established by the National Center for Health Statistics. More detail is available in the separate Cause of Death dataset.	Character " "="Missing"
death_days	Days since randomization at death from most definitive source	Days from randomization to death.	Numeric .N="No date of death on record"
deathcutoff	Death included in lung cancer mortality analysis	Indicates whether deaths occurred before the cutoff for the official final analysis of lung cancer mortality (January 15, 2009). Note that the database was updated after the final analysis, so the data differ slightly from the published results of the trial.	0="No death or no date of death" 1="Death Included" 2="Death Not Included"
deathstat	Death status	Indicates completeness of the death confirmation process.	0="No report of death" 1="EVP certified" 2="Death Certificate coded" 3="Death Certificate received but not coded" 4="Death reported, DC expected to be obtained" 5="Death reported, DC cannot be obtained" 6="NDI Exact match, DC not obtained" 7="NDI Probable match, DC not obtained"
finaldeathlc	Final lung cancer death (Combined best information EVP supplemented with DCF)	This is the authoritative variable indicating whether lung cancer was the cause of death. It uses the information from the Endpoint Verification Process (EVP) if available and uses the death certificate if EVP is not available.	.M="Death Reported, cause of death unknown" .N="No death reported" 0="Death not due to lung cancer" 1="Death due to lung cancer or work-up of suspected lung cancer"

Variable	Label	Description	Format Text
hasdcf	Death certificate status	Has a death certificate been received? Some participants went through the Endpoint Verification Process (EVP), and their lung cancer death status was certified, without a death certificate being received.	0="No" 1="Yes"
ndicd	ICD-10 code for underlying cause of death as provided from NDI+ search.	The cause of death from a National Death Index search. This variable is of little value for analysis. It is available for very few participants and is not as reliable as information from the death certificate (in DCFICD). This variable should only be populated if deathstat is 6 or 7.	Character " "="Missing"

Section 18: Coronary Artery Calcification

Variable	Label	Description	Format Text
cac_case0-2	Case/Control status for the participant in study year TX for the CAC selection.	This variable contains the case/control status for the participant in study year TX for the Coronary Artery Calcification (CAC) selection. CAC is categorized based on string searches of the verbatim text using keywords associated with CAC and the pattern of CAC responses across study years. In three study centers CAC was thoroughly annotated in the radiologist reports; in six of the study centers CAC was anecdotally reported; and the remainder of the study centers did not record CAC at all.	.A="LSS 7 Centers" .N="ACRIN, XRY or No Screen" 0="CAC Not Detected" 1="CAC Detected"
cac_case_ever	Case/Control status for the participant in at least one Study Year for the CAC selection.	Contains the case/control status for the participant across all study years for the Coronary Artery Calcification (CAC) selection. If the participant was ever a case in T0, T1, or T2, they are considered a case. CAC is categorized based on string searches of the verbatim text using keywords associated with CAC and the pattern of CAC responses across study years. In three study centers CAC was thoroughly annotated in the radiologist reports; in six of the study centers CAC was anecdotally reported; and the remainder of the study centers did not record CAC at all.	.A="LSS 7 Centers" .N="ACRIN, XRY or No Screen" 0="CAC Not Detected" 1="CAC Detected"

Variable	Label	Description	Format Text
cac_case_num	Number of Study Years for which the participant is a case for the CAC selection.	Number of study years in which Coronary Artery Calcification (CAC) was detected among participants in the CAC selection. CAC is categorized based on string searches of the verbatim text using keywords associated with CAC and the pattern of CAC responses across study years. In three study centers CAC was thoroughly annotated in the radiologist reports; in six of the study centers CAC was anecdotally reported; and the remainder of the study centers did not record CAC at all.	Numeric .A="LSS 7 Centers" .N="ACRIN, XRY or No Screen"
selected_cac_pop	Selected for Coronary Artery Calcification (CAC) Image Population	Denotes whether a participant was selected for the Coronary Artery Calcification selection. Eligible participants were from the 3 Centers (Marshfield, Georgetown, Henry Ford), with at least one study year with an axial image.	0="No" 1="Yes"
selected_cac_pop_batch	Coronary Artery Calcification Batch	Coronary Artery Calcification (CAC) batch number. Participants in the CAC selection were divided into two evenly-distributed sets which were further each divided into 6 batches of comparable size.	Numeric .N="ACRIN, XRY or No Screen"
selected_cac_pop_case	Coronary Artery Calcification Case/Control	Case/Control status of participants selected for the Coronary Artery Calcification (CAC) selection. Cases are comprised of participants with mentions of CAC from any source in at least one study year, and no mention of CAC in the other study years. Controls are comprised of participants with no mention of CAC for any study year.	.N="ACRIN, XRY or No Screen" 1="Case" 2="Control"

Variable	Label	Description	Format Text
selected_cac_pop_casestat	CAC Case/Control Status	Detailed Case/Control status of participants selected for the Coronary Artery Calcification (CAC) selection. Cases are comprised of participants with mentions of CAC from any source in at least one study year, and no mention of CAC in the other study years. Controls are comprised of participants with no mention of CAC for any study year.	.C="X-ray arm participant" 1="Yes in all SY from 3 centers with at least 2 screens" 2="All No's then All Yes's from 3 centers with at least 2 screens" 3="No following Yes from 3 centers" 4="Single yes from 3 centers, only 1 screen" 5="Yes in all SY from 6 centers with at least 2 screens" 6="All No's then All Yes's from 6 centers with at least 2 screens" 7="No following Yes from 6 centers" 8="Single yes from 6 centers, only 1 screen" 11="Always No from 3 centers and had 3 screens" 12="Always No from 3 centers and < 3 screens" 13="Always No from 6 centers and had 3 screens" 14="Always No from 6 centers and had < 3 screens" 15="Never Screened" 16="Colorado" 17="ACRIN participant"
selected_cac_pop_casestat_simple	CAC Case/Control Status (simplified)	Summary of Case/Control status based on the variable selected_cac_pop_casestat. CAC is categorized based on string searches of the verbatim text using keywords associated with CAC and the pattern of CAC responses across study years. In three study centers CAC was thoroughly annotated in the radiologist reports; in six of the study centers CAC was anecdotally reported; and the remainder of the study centers did not record CAC at all.	.C="X-Ray Arm" 1="3 Centers Case" 2="3 Centers Control" 3="6 Centers Case" 4="6 Centers Control" 5="Colorado / ACRIN / Never Screened"
selected_cac_pop_set	Coronary Artery Calcification Set	Coronary Artery Calcification (CAC) set number. Participants in the CAC selection were divided into two evenly-distributed sets designed to be used for training/testing purposes.	.N="ACRIN, XRY or No Screen" 1="Set 1" 2="Set 2"

Section 19: Trial Endpoint Verification Process (EVP)

Variable	Label	Description	Format Text
evp_revr	Highest level of EVP completed	What was the highest level of the Endpoint Verification Process (EVP) completed? "CDQ" stands for Cause of Death Questionnaire, which is the form completed by EVP team members.	.N="No death certificate coded (or not deceased)" 0="No chair-level CDQ completed (includes if only CDQ received says chair unblinded or needs more information)" 1="Chair-level CDQ completed (includes if chair was unblinded and a member completed the chair-level review)" 2="Member-level CDQ completed" 3="Team CDQ completed" 4="No CDQ completed , no records available"
evpcert	EVP Certified or not	Was the death certified by the Endpoint Verification Process (EVP), with or without review?	.N="No death certificate coded (or not dead)" 0="Not Certified" 1="Certified (with or without review)"
evpdeath	EVP Certified cause of death	Cause of death according to the Endpoint Verification Process (EVP).	.N="Not Certified or Not Selected" 1="Death due to lung cancer" 2="Other cancer" 3="Not cancer" 4="Death due to diagnostic evaluation for a suspected lung cancer"
evpdirect	Is death the direct, indirect, or direct and indirect result of lung cancer?	For EVP certified deaths from lung cancer (evpdeath is 1 or 4), is death the direct, indirect, or direct and indirect result of lung cancer?	.N="Death NOT due to lung cancer or no death (EVPDEATH in 2, 3, .N)" 1="Direct, result of lung cancer" 2="Indirect, result of diagnostic evaluation" 3="Indirect, result of treatment" 5="Direct and indirect, result of lung cancer and diagnostic evaluation" 6="Direct and indirect, result of lung cancer and treatment" 8="None of the above/Incomplete"
evpincomplete	Reason why a CDQ/EVP review could not be completed	Reason why the Endpoint Verification Process (EVP) review could not be completed.	.N="Not Applicable" 5="Medical records do not exist" 6="Medical records cannot be obtained" 7="Medical records are inadequate and additional documentation cannot be obtained" 8="No resolution was reached before the study was locked."
evpsel	EVP selected or not	Was the death selected for review by the Endpoint Verification Process (EVP)?	.N="No death certificate coded (or not dead)" 0="Run through algorithm and not selected for EVP" 1="Selected for EVP" 2="Late algorithm run (near study closeout date); not selected for EVP"
evpsent	Sent to EVP or not	Did the death undergo a review by the Endpoint Verification Process (EVP)?	.N="No records available" 0="Records not sent to EVP" 1="Records sent to EVP"