

NLST CT Image Info Masterfile: Data Dictionary

TABLE OF CONTENTS

Document Summary.....	2
NLST CT Image Info Masterfile: Data Dictionary	3
Section 1: Study	3
Section 2: Overview	4
Section 3: Reconstruction Info.....	7
Section 4: Scanner Info	8
Section 5: Technical Parameters	9

Document Summary

Property	Value
Document Title	NLST CT Image Info Masterfile: Data Dictionary
Date Created	02/17/2026
Sections	5
Entries	38
Document Filename	dictionary_sctimageinfo-20110203-021326.rtf

NLST CT Image Info Masterfile: Data Dictionary

Section 1: Study

Variable	Label	Description	Format Text
dataset_version	Date Stamp for Datasets		Char, 23
pid	Participant identifier, also called the IMS Global ID	LSS Participants (1xxxxx) ACRIN Participants (2xxxxx)	Numeric
seriesinstanceuid	New series instance uid	Uniquely identifies each image series. For each screening round (T0, T1, T2), most participants have multiple image series.	Char
study_yr	Study year of screening exam		0="T0" 1="T1" 2="T2"
studyuid	New study instance uid	Unique identifier for each CT study (i.e. screening round). For all series from a given screening round (T0, T1, or T2), each participant has a single value of studyuid.	Char

Section 2: Overview

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
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.	1="Set 1" 2="Set 2"
imageclass	Image class (0/1 value for localizer/axial images)		0="Localizer" 1="Axial" 9="Undetermined"
imagesize_kilobytes	Total file size of images in this series (in kilobytes)		Numeric
imagetype	Image type (localizer or axial images)	If the text contains "axial", then this image series is from an axial reconstruction of a detailed full chest CT scan with many slices. If the text contains "localizer", or "top", then the series consists of one or two scout images of the entire chest, performed prior to the detailed CT scan to identify the endpoints for the full scan (just above lung apices, just below costophrenic angles).	Char
numberimages	Number of images contained within a series		Numeric
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"

Variable	Label	Description	Format Text
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_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"
selected_lc_pop	Image series is included in 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.	0="No" 1="Yes"
selected_lc_pop_batch	SCT Image Batch For Lung Cancer Selection	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.	.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"

Variable	Label	Description	Format Text
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.	.N="Not Applicable" 1="Set 1" 2="Set 2"
seriesdescription	A description of the series as a list* of values found within or computed from values in the DICOM image headers.	* Series Description: a concatenation of 11 comma-separated values (exact or coded) representing: (1) Screening year (2) Image Type (3) Manufacturer (4) Model (5) Convolution Kernel (6) Reconstruction Diameter (7) Slice Thickness (8) kVp (9) mAs (10) Effective mAs (11) Pitch	Char

Section 3: Reconstruction Info

Variable	Label	Description	Format Text
filtertype	Filter Type		Char " "="Missing"
reconfilter	Name of the reconstruction convolution kernel used by the manufacturer to render the spiral computed tomography (SCT) images.		Char " "="Missing"
reconstruction_diameter	The millimeter diameter of the region from which the data were used in creating the reconstruction of the image.		Numeric .M="Missing"
reconstruction_diameter_raw	Raw character value for reconstruction diameter.	0 - 1750	Char " "="Missing"
reconthickness	Reconstruction slice thickness		Numeric .M="Missing"
reconthickness_raw	Raw character value for reconstruction slice thickness.	0.75 - 999	Char " "="Missing"

Section 4: Scanner Info

Variable	Label	Description	Format Text
manufacturer	Name of the manufacturer of the spiral computed tomography scanner.		1="GE Medical Systems" 2="Philips" 3="Siemens" 4="Toshiba"
manufacturer_raw	Raw character value for SCT scanner manufacturer.	GE Medical Systems, Phillips, Siemens, Toshiba	Char
manufacturers_model_name	Model name given to the scanner by the manufacturer.		Char
softwareversion	Scanner software version at the time of the scan.		Char " " = "Missing"

Section 5: Technical Parameters

Variable	Label	Description	Format Text
effmas	Effective mAs		Numeric .M="Missing"
exposure_time	Exposure Time		Numeric .M="Missing"
exposure_time_raw	Raw character value for exposure time.		Char " "="Missing"
kvp	Kilo Volt peak (kVp) - peak kilo voltage output of the x-ray generator used.		Numeric .M="Missing"
kvp_raw	Raw character value for kilo-volt peak.	80, 90, 100, 110, 120, 130, 135, 140	Char
mas	Milli-ampere second (mAs) - Exposure expressed as milli-ampere seconds calculated from exposure time and x-ray tube current.		Numeric .M="Missing"
pitch	Pitch is the table travel per rotation divided by the collimation.		Numeric .M="Missing"
pitch_raw	Raw character value for pitch.	0.75, 0.875, 1.25, 1.375, 1.5, 1.75	Char " "="Missing"
xray_tube_current	X-ray Tube Current		Numeric .M="Missing"
xray_tube_current_raw	Raw character value for x-ray tube current.		Char " "="Missing"