

Lung X-Ray Images 25K (lung_xry_images_25k) Data Dictionary

TABLE OF CONTENTS

Document Summary.....	2
Lung X-Ray Images 25K (lung_xry_images_25k): Data Dictionary.....	3
Section 1: Identifiers	3
Section 2: X-Ray Images	4

Document Summary

Property	Value
Document Title	Lung X-Ray Images 25K (lung_xry_images_25k): Data Dictionary
Date Created	10/15/2024
Sections	2
Entries	10
Document Filename	dictionary_lung_xry_images_25k-t20241011.rtf

Lung X-Ray Images 25K (lung_xry_images_25k): Data Dictionary

Section 1: Identifiers

Variable	Label	Description	Format Text
build	Build	Masterfile build. Distributed to all datasets and used to identify the version of the database.	Char, 30
build_cancers	Build Cancer Cohort	The source of the build data.	1="Trial Only"
build_incidence_cutoff	Build Incidence Cutoff	The censoring date for any cancer data associated with the dataset.	1="Cancers Through 2009"
plco_id	PLCO ID		Char

Section 2: X-Ray Images

Variable	Label	Description	Format Text
in_ccsub_selection	Image Is In The 2020 Case/Control Subset	Image was selected as a case or control in the case/control subset of the 25k selection. The Case/Control Subset is a selection of x-ray images designed for the analysis of lung cancer that is available on CDAS as a manageable download. The subset was constructed with screen-detected and interval cancers matched to cancer-free controls.	0="No" 1="Yes"
selected_25k_2020_batch	Download Batch For Standard 25k Images	Image download batch associated with images in Standard 25k Selection. The Standard 25k Selection is a selection of 25,000 PLCO participants with X-ray images that is available for download from CDAS. It includes all participants with T0-T5 lung cancer or non-target cancer, all participants with any abnormal/suspicious screens, a random selection of participants with any abnormal/not suspicious screens, and a random selection of participants with all negative screens.	1="Batch A" 2="Batch B" 3="Batch C" 4="Batch D" 5="Batch E" 6="Batch F"
selected_25k_2020_set	Set For Standard 25k Selection Images	Image download set associated with images in Standard 25k Selection. The Standard 25k Selection was divided into two equally-distributed sets that were intended to be used as training and test sets. The Standard 25k Selection is a selection of 25,000 PLCO participants with X-ray images that is available for download from CDAS. It includes all participants with T0-T5 lung cancer or non-target cancer, all participants with any abnormal/suspicious screens, a random selection of participants with any abnormal/not suspicious screens, and a random selection of participants with all negative screens.	1="Set 1" 2="Set 2"
study_yr	Study Year of Linked Screen	Screen year for images positively linked to a screen	Numeric
image_file_name	Name of Image on Disk		Char
image_file_size	Image File Size in Bytes		Numeric

