

TCIA NLST-LSS Pathology Dataset: Data Dictionary

Table of Contents 8/29/2013

Section Number	Section Title
1	Identifiers
2	Tumor Information
3	Tumor-wide TMA (tissue micro array)/ DNA Core Information
4	Block Region Information
5	Block TMA Cores
6	Region Summary
7	Region Cancer Variables
8	Region Non-Tumor Histologies
9	Region Pre-Malignant Histologies
10	TMAcore Sample Information

TCIA NLST-LSS Pathology Dataset: Data Dictionary

Section 1: Identifiers

Class	Variable	Label	Description	Format Text
1. Principal	alpha_beta	In Alpha or Beta TMA block?	Is the TMA tissue core from the Alpha or Beta block? For each TMA layout (as distinguished by the TMA_BLOCK variable), there were two duplicate blocks constructed, the Alpha and the Beta blocks.	1="Alpha" 2="Beta"
1. Principal	co_slide_id	Colorado Slide Identifier for the Pathology Images	Identifier for images of pathology slides. As assigned by annotating path lab and transcribed by TMA construction path lab.	char, 11
1. Principal	lc_order	Order of this tumor among all tumors diagnosed	For participants with multiple primary lung cancers, this variable indicates which cancer is being described.	1="1st lung cancer"
1. Principal	pathology_image	Pathology Image ID (Same as tissue_block_id)	This identifier is used to create a ViewPathologyImage button to allow viewing of the pathology images with an image viewer. Format: AA##### ####	char, 15 null ="Missing"
1. Principal	pid	Participant Identifier		Numeric
1. Principal	roi_number	Number of ROI on donor block	Each Region of Interest within a tissue block is assigned a number. This is used for linking TMA cores to the roi.	Numeric
1. Principal	study_yr	Study year associated with first confirmed lung cancer		Numeric
1. Principal	tissue_block_id	Identifier of donor tissue block	ID of original Tissue Block	char, 12
1. Principal	tma_block	TMA Block Name	Label for each tissue microarray block. Note that for each letter, two TMA blocks (with identical layouts) exist. These duplicate blocks are the Alpha and Beta blocks, distinguishable by using the alpha_beta variable.	"A NEO"="A NEO" "A"="A" "B"="B" "C"="C" "D"="D" "E"="E" "F"="F" "G"="G" "H"="H"
1. Principal	tma_block_id	Identifier of TMA block in which core is arrayed	Format: AA##### ####	char, 14
1. Principal	tmacore_id	TMA Core Identifier	Format: AA##### ####	char, 14
1. Principal	tumor_ID	Tumor identifier	ID of original tumor	char, 7

Section 2: Tumor Information

Class	Variable	Label	Description	Format Text
1. Principal	case_highest_inv_grade	Highest grade observed in tumor		0="Blank" 1="Well Differentiated" 2="Moderately" 3="Poorly" 4="Undifferentiated" 99="Missing"
1. Principal	case_likely_metastases	Case likely to have metastases?		0="None" 1="Probable" 2="Unlikely" 9="Cannot Determine"
1. Principal	case_longest_inv_diam	Longest diameter of invasive component of tumor	From central pathology review. Sometimes multiple entries are entered, comma-delimited. Two entries implies two WHO Classes given, but not vice-versa.	char, 8
1. Principal	case_side	Side of body where tumor is located	Side of body on which tumor is located. From original path report.	1="Right" 2="Left"
1. Principal	case_topography_descrip	Topography Description	Character Field for Tumor Location(s) (from original path report)	Char, 64
1. Principal	case_who_class	WHO classification of tumor	Raw WHO Classification (2004) in Character Format. Usually 6 characters, #####/# Occasionally contains 2 separate classifications for the same tumor.	char, 14

Section 3: Tumor-wide TMA (tissue micro array) / DNA Core Information

Class	Variable	Label	Description	Format Text
1. Principal	case_any_tmacore	Any tissue from this tumor in the tissue microarrays?	Indicator for TMA Cores	0="No" 1="Yes"
1. Principal	case_tmacore_cnt	Number of tissue cores from this tumor in the tissue microarrays	TMA Core Count	Numeric

Section 4: Block Region Information

Class	Variable	Label	Description	Format Text
1. Principal	block_annotated	Were ROIs annotated on the block?	Were regions of interest annotated for this block? By protocol, digital annotations were made on an image of an H&E slide of each tissue block. Each annotation color was used for a specific histology.	0="No" 1="Yes"

Section 5: Block TMA Cores

Class	Variable	Label	Description	Format Text
1. Principal	block_has_tmacore	Any tissue from this donor block in the TMAs?	Indicator for whether or not Tissue block has TMA cores or not.	0="No" 1="Yes"
9. Secondary	block_tmacore_orig_cnt	Count of cores in the TMAs from this block	Number of TMA Cores before any were used	Numeric

Section 6: Region Summary

Class	Variable	Label	Description	Format Text
1. Principal	roi_features_comments	Misc. comments about ROI	Miscellaneous comments	Char, 107 null ="Missing"
1. Principal	roi_histology_cat	Histology category (=> annotation color)		1="Metastasis" 2="Invasive Carcinoma" 3="Carcinoma In Situ" 4="Pre-Malignant" 5="Distal Bronchiole" 6="Proximal Bronchus" 7="Normal Lymph Node" 8="Normal Lung" 9="Other"
1. Principal	roi_histology_subtype	Histology: Text description	For cancers, it includes a text description of the WHO classification. For non-tumor lung tissue, see roi_histology_cat for further distinctions (i.e. distal bronchiole vs. proximal bronchus vs. other normal lung tissue).	Char, 65
1. Principal	roi_label_color	Annotation color (=> histology)	Color used by pathologist to mark the regions of interest of each tissue block. Each color corresponds to a particular histology (see roi_histology_cat).	"AQUA"="Aqua (Metastasis)" "BLACK"="Black (Invasive Carcinoma)" "BLUE"="Blue (Carcinoma In Situ)" "BROWN"="Brown (Pre-Malignant)" "GOLD"="Gold (Other)" "GREEN"="Green (Distal Bronchiole)" "ORANGE"="Orange (Normal Lymph Node)" "OTHER"="Other" "PURPLE"="Purple (Normal Lung)" "RED"="Red (Invasive Type II)" "YELLOW"="Yellow (Proximal Bronchus)"

Section 7: Region Cancer Variables

Class	Variable	Label	Description	Format Text
1. Principal	roi_highest_grade	Highest grade observed in ROI		0="Blank" 1="Well Differentiated" 2="Moderately" 3="Poorly" 4="Undifferentiated"
1. Principal	roi_inflamm_cell_percent	Percent inflammatory cells within ROI		Numeric null ="Missing"
1. Principal	roi_inv_tumor_dim	Size of Invasive Component of Tumor (mm)	In millimeters	Numeric null ="Missing"
1. Principal	roi_lymph_inv	Lymphatic invasion in ROI?		Null ="Missing" Not Applicable 0="No" 1="Yes"
1. Principal	roi_metastases	Metastases: likely from tumor tissue in this ROI?		Missing None Not Applicable 0="No" 1="Probable" 2="Unlikely" 99="Can't Determine"
1. Principal	roi_tumor_cell_percent	Percent tumor cells within ROI		Numeric null ="Missing"
1. Principal	roi_vessel_inv	Blood vessel invasion in ROI?		Null ="Missing" Not Applicable 0="No" 1="Yes"
1. Principal	roi_who_class	WHO classification of tumor tissue in ROI	Based on 2004 WHO classification of lung tumors. Includes a 4-digit morphology code and a 1-digit behavior code (####/#).	Char, 6 null ="Missing"
1. Principal	roi_who_class_beh	ICD-O-3 behavior code from WHO classification		null ="Missing" 2="In-Situ" 3="Invasive"
1. Principal	roi_who_class_mor	ICD-O-3 morphology code from WHO classification	Whole number with 4 digits.	Numeric null ="Missing"

Section 8: Region Non-Tumor Histologies

Class	Variable	Label	Description	Format Text
1. Principal	roi_nthist_chronic_bronchitis	Has Chronic bronchitis in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_emphysema	Has Emphysema in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_fibrosis	Has Fibrosis in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_granuloma	Has Granuloma in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_hemorrhage	Has Hemorrhage in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_infarction	Has Infarction in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_necrosis	Has Necrosis in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_normal_lung	Has Normal lung tissue in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_other_specify	Has Other (SPECIFY) non-tumor histology in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_other_specify_text	Has Other Condition - Verbatim text		Char, 57 null ="Missing"
1. Principal	roi_nthist_pneumonia	Has Pneumonia in ROI		0="No" 1="Yes"
1. Principal	roi_nthist_pre_neoplastic_tiss	Has Pre Neoplastic tissue in ROI		0="No" 1="Yes"

Section 9: Region Pre-Malignant Histologies

Class	Variable	Label	Description	Format Text
1. Principal	roi_nontumor_premalig	Pre-malignant tissue type		1="Squamous Carcinoma in-Situ" 2="Squamous Dysplasia, MILD" 3="Squamous Dysplasia, Moderate" 4="Squamous Dysplasia, SEVERE" 5="Atypical Adenomatous Hyperplasia (AAH)" 7="Reserve Cell Hyperplasia" 8="Carcinoid Tumorlet" 88="Other" 99="N/A"

Section 10: TMAcore Sample Information

Class	Variable	Label	Description	Format Text
1. Principal	micro_array_col	Micro Array column	= 2 to 15 for quadrants 1 & 2, 22 to 35 for quadrants 3 & 4. Columns 1 & 21 are orientation columns containing non-NLST tissue.	numeric
1. Principal	micro_array_row	Micro Array row	= 1 to 9 for quadrants 1 & 3, 21 to 29 for quadrants 2 & 4	numeric
1. Principal	quadrant	Micro Array quadrant		1="Upper Left" 2="Lower Left" 3="Upper Right" 4="Lower Right"