Memorial Sloan-Kettering Lung Study: Summary Person Data Dictionary

TABLE OF CONTENTS

Document Summary

Property	Value
Document Title	Memorial Sloan-Kettering Lung Study: Summary Person: Data Dictionary
Date Created	04/16/2024
Sections	10
Entries	96
Document Filename	dictionary_mem-external_summary_person-04102024.rtf

Memorial Sloan-Kettering Lung Study: Summary Person: Data Dictionary

Section 1: Identifiers

Class	Variable	Label	Description	Format Text
01. Principal	**pid**	Patient ID	Unique Participant ID	Char, 7

Section 2: Trial Entry/Exit

Class	Variable	Label	Description	Format Text
01. Principal	**arm**	Arm	Randomization Arm	1="Dual Screen Arm" 2="Single Screen Arm"
01. Principal	**basic_age_at_randomization**	Age at Randomization		Numeric
01. Principal	**dx_to_mortality_exitdays**	Days from Diagnosis to Death or Last Follow-Up	This variable exists only in the Summary file.	Numeric .N="Not Applicable"
01. Principal	**dx_to_mortality_exitmonths**	Months from Diagnosis to Death or Last Follow-Up	This variable exists only in the Summary file.	Numeric .N="Not Applicable"
01. Principal	**entry_to_dx_days**	Days from Entry to Diagnosis	This variable exists only in the Summary file.	Numeric .N="Not Applicable"
01. Principal	**entry_to_dx_months**	Months from Entry to Diagnosis	This variable exists only in the Summary file.	Numeric .N="Not Applicable"
01. Principal	**entry_to_mortality_exitdays**	Days from Entry to Death or Last Follow-Up	This variable exists only in the Summary file.	Numeric .M="Missing Value"

Class	Variable	Label	Description	Format Text
01. Principal	**entry_to_mortality_exitdays_v2**	Interval 1 (Days)	This variable exists only in the Summary file. Revised intervals from entry to death: If the patient is not dead and is not lost to follow-up all three time intervals are the same as the interval from entry to death. If the patient is not dead and the year of the calculated date of death or last follow-up is before 1982, then the patient is lost to follow-up and the intervals are calculated in the following ways: a. Interval 1: If the patient is in the screened group then Interval 1 = date last - date of entry; otherwise, interval 1 is the same as the interval from entry to death. Date last is the date of the last radiology record that has a month and day that falls within 2 months on either side of the month and day of the patient's entry date.	Numeric .M="Missing Value"
01. Principal	**entry_to_mortality_exitdays_v3**	Interval 2 (Days)	This variable exists only in the Summary file. Revised intervals from entry to death: If the patient is not dead and is not lost to follow-up all three time intervals are the same as the interval from entry to death. If the patient is not dead and the year of the calculated date of death or last follow-up is before 1982, then the patient is lost to follow-up and the intervals are calculated in the following ways: b. Interval 2 = interval 1 + 183 days (1/2 year)	Numeric .M="Missing Value"

Class	Variable	Label	Description	Format Text
01. Principal	**entry_to_mortality_exitdays_v4**	Interval 3 (Days)	This variable exists only in the Summary file. Revised intervals from entry to death: If the patient is not dead and is not lost to follow-up all three time intervals are the same as the interval from entry to death. If the patient is not dead and the year of the calculated date of death or last follow-up is before 1982, then the patient is lost to follow-up and the intervals are calculated in the following ways: c. Interval 3 = 12/31/83 - date of entry	Numeric .M="Missing Value"
01. Principal	**entry_to_mortality_exitmonths**	Months from Entry to Death or Last Follow-Up	This variable exists only in the Summary file.	Numeric .M="Missing Value"
11. Process	**basic_randomization_dt_adj**	Adjusted Date of Randomization?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	0="No" 1="Yes"

Section 3: Baseline Characteristics

Class	Variable	Label	Description	Format Text
01. Principal	basic_amount_of_cig_smoked	Amount of Cigarette Smoked		1="Less Than 1/2" 2="1/2 to 3/4" 3="Greater Than 3/4"
01. Principal	basic_cancer_epidermis	Personal History of Cancer: Epidermoid Neoplasm		0="No" 1="Yes"
01. Principal	basic_cancer_hematopoietic	Personal History of Cancer: Malignant Neoplasm Hematopoietic Tissue	This personal history of cancer variable is of questionable quality, but was retained for the sake of completeness.	0="No" 1="Yes"
01. Principal	basic_cancer_lymphoma	Personal History of Cancer: Lymphoma		0="No" 1="Yes"
01. Principal	basic_cancer_thyroid	Personal History of Cancer: Thyroid	This personal history of cancer variable is of questionable quality, but was retained for the sake of completeness.	0="No" 1="Yes"
01. Principal	basic_cancer_unspecified	Personal History of Cancer: Unspecified	This personal history of cancer variable is of questionable quality, but was retained for the sake of completeness.	0="No" 1="Yes"
01. Principal	basic_cigpd_year_ago	Cigarettes per Day (1 year ago)		Numeric
01. Principal	basic_race	Race		1="White" 2="Black" 3="Spanish" 4="Other"
01. Principal	basic_sex	Sex		1="Male" 2="Female"
01. Principal	basic_smoke_inhale	Inhale?	Indicates whether the participant inhaled cigarette smoke.	0="No" 1="Yes"
01. Principal	basic_smoke_inhale_deeply	Inhale Deeply?	Indicates whether the participant inhaled cigarette smoke deeply.	0="No" 1="Yes"
01. Principal	basic_smoke_month_stopped	Month Stopped Smoking		Numeric 0="Still Smoking"
01. Principal	basic_smoke_stopped	Stopped Smoking?		0="No" 1="Yes"
01. Principal	basic_smoke_yr	Years Smoked		Numeric

Section 4: Radiology

Class	Variable	Label	Description	Format Text
01. Principal	**rad_counter**	Number of Entries for Radiology		Numeric .F="No Form"
01. Principal	**rad_days1-22**	Days from Randomization to Date of Radiology Reading		Num .F="No Form" .N="Not Applicable"
01. Principal	**rad_study_radiol ogist_dx1-22**	Diagnosis of Study Radiologist		.F="No Form" .N="Not Applicable" 0="No Response" 1="Negative" 2="Suspicious" 3="Cancer" 4="Unsatisfactory" 9="Invalid Response"
01. Principal	**rad_study_radiol ogist_rec1-22**	Recommendation of Study Radiologist		.F="No Form" .N="Not Applicable" 0="No Response" 1="Submitted Films" 2="Repeat X-Ray" 3="Consultation" 4="4-Month Repeat" 5="M.D. Referral" 9="Invalid Response"
11. Process	**rad_dt_adj1-22**	Adjusted Date of Radiology Reading?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.F="No Form" .N="Not Applicable" 0="No" 1="Yes"

Section 5: Cytology

Class	Variable	Label	Description	Format Text
01. Principal	**cyto_can1-44**	Cytology Interpretation: Cancer		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
01. Principal	**cyto_counter**	Number of Entries for Cytology		Numeric .F="No Form"
01. Principal	**cyto_mod1-44**	Cytology Interpretation: Moderate		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
01. Principal	**cyto_mos_since_ init_visit1-44**	Months Since Initial Visit (Values Are Multiples Of 4)		Numeric .F="No Form" .N="Not Applicable"
01. Principal	**cyto_neg1-44**	Cytology Interpretation: Negative		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
01. Principal	**cyto_other1-44**	Cytology Interpretation: Other		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
01. Principal	**cyto_recommend ation1-44**	Cytology Recommendation		.F="No Form" .N="Not Applicable" 0="No Response" 1="Routine Follow-Up" 2="Repeat Spontaneous" 3="Repeat Aerosol" 4="Thoracic Consultation" 9="Invalid Response"
01. Principal	**cyto_sli1-44**	Cytology Interpretation: Slight		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
01. Principal	**cyto_sus1-44**	Cytology Interpretation: Suspicious		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
01. Principal	**cyto_uns1-44**	Cytology Interpretation: UNS		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"

Section 6: Cancer Diagnosis

Class	Variable	Label	Description	Format Text
01. Principal	**cncr_det_confirmation**	Detection Confirmation	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Resected Specimen" 2="Thoracotomy and Biopsy" 3="Other Biopsy" 4="Bronchial Biopsy" 5="Needle Aspiration" 6="Radiological and Clinical" 7="Bronchial Cytology" 8="Effusion Cytology" 9="Sputum Cytology" 10="Cerebral Cytology" 11="Autopsy"
01. Principal	**cncr_det_status**	Detection Status	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Confirmed" 2="Probable"
01. Principal	**cncr_detection_days**	Days from Randomization to Date of Detection	Populated only for participants with probable or confirmed lung cancer.	Numeric .F="No Form"
01. Principal	**cncr_detection_method**	Method of Detection	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Routine Cytology" 2="Routine X-ray" 3="X-ray and Cytology" 4="Interval" 5="Post-Screening"
01. Principal	**cncr_measure_method**	Method of Measurement	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Resected Specimen" 2="Surgery - No Resection" 3="X-Ray" 4="No Measurement" 5="A (Undocumented Code)" 6="B (Undocumented Code)" 7="U (Undocumented Code)"
01. Principal	**cncr_preval_incidence**	Prevalence/Incidence	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Incidence" 2="Prevalence"
01. Principal	**cncr_staging_basis**	Basis for Staging	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Pathological" 2="Clinical" 3="Surgical" 4="Unstaged" 5="Autopsy"
01. Principal	**dx_age**	Age at Diagnosis (Years)	Populated only for participants with probable or confirmed lung cancer.	Numeric .N="Not Applicable"

Class	Variable	Label	Description	Format Text
01. Principal	**enrollment_age**	Age at Enrollment	Populated only for participants with probable or confirmed lung cancer.	Numeric .F="No Form"
01. Principal	**lung_cancer_flag**	Lung Cancer Flag		0="No Lung Cancer" 1="Confirmed Lung Cancer"
09. Secondary	**cncr_det_aerosol _result**	Aerosol Sputum at Detection: Result	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 0="Negative" 1="Slight" 2="Moderate" 3="Cancer" 4="Unsatisfactory" 9="S (Undocumented Code)"
09. Secondary	**cncr_det_aerosol _type**	Aerosol Sputum at Detection: Type	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 0="Not Done" 1="Routine" 3="Special"
09. Secondary	**cncr_det_spont_r esult**	Spontaneous Sputum at Detection: Result	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 0="Negative" 1="Slight" 2="Moderate" 3="Cancer" 4="Unsatisfactory" 9="S (Undocumented Code)"
09. Secondary	**cncr_det_spont_t ype**	Spontaneous Sputum at Detection: Type	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 0="Not Done" 1="Routine" 2="Hospital" 3="Special"
09. Secondary	**cncr_det_xray_re sult**	X-ray at Detection: Result	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 0="Negative" 1="Suspicious" 2="Cancer"
09. Secondary	**cncr_det_xray_ty pe**	X-ray at Detection: Type	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 0="Not Done" 1="Routine" 2="Hospital" 3="Special"
11. Process	**cncr_detection_d t_adj**	Adjusted Date of Detection?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.F="No Form" 0="No"

Section 7: Cancer Characteristics

Class	Variable	Label	Description	Format Text
01. Principal	**cncr_ajcc_stage**	Stage	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 0="Stage 0" 1="Stage I" 2="Stage II" 3="Stage III" 8="Unknown"
01. Principal	**cncr_cell_type**	Cell Type	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Adenocarcinoma" 2="Epidermoid (Squamous)" 3="Oat Cell" 4="Large Cell" 5="Carcinoid" 8="Unknown"
01. Principal	**cncr_recurrence_days**	Days from Randomization to Date of Recurrence (for resected cases only)	Populated only for participants with probable or confirmed lung cancer.	Numeric .F="No Form" .M="Missing Value"
01. Principal	**cncr_recurrence_site**	Site of Recurrence	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 1="Brain and Bone" 2="Brain" 3="Effusion" 4="Liver" 5="Lung Bilateral" 6="Lung Other" 7="Lung Same" 8="Multiple Mets" 9="Node" 10="Bone" 11="Primary Lung"
01. Principal	**cncr_resected_nodes**	Nodes Resected	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 0="None" 1="All" 2="Hilar" 3="Mediastinal" 4="Peribronchial"
01. Principal	**last_contact_days**	Days from Randomization to Date of Last Contact	Populated only for participants with probable or confirmed lung cancer.	Numeric .F="No Form"
01. Principal	**mortality_review_days**	Days from Randomization to Date of Mortality Review	Populated only for participants with probable or confirmed lung cancer.	Numeric .F="No Form" .M="Missing Value"
01. Principal	**status_final**	Final Status	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Alive" 2="Dead"

Class	Variable	Label	Description	Format Text
03. Components	**cncr_ajcc_metast ases**	M Component of TNM Staging	The M-Component of the TNM Stage. Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 0="M0" 1="M1" 8="Unknown"
03. Components	**cncr_ajcc_node_ stage**	N Component of TNM Staging	The N-Component of the TNM Stage. Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 0="N0" 1="N1" 2="N2" 8="Unknown"
03. Components	**cncr_tumor_stag e**	T Component of TNM Staging	The T-Component of the TNM Stage. Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 1="T1" 2="T2" 3="T3" 4="T: In Situ" 5="T: Microinvasive" 8="Unknown"
11. Process	**cncr_recurrence_ dt_adj**	Adjusted Date of Recurrence (for resected cases only)?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.F="No Form" .M="Missing Value" 0="No"
11. Process	**last_contact_dt_ adj**	Adjusted Date of Last Contact?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.F="No Form" 0="No"

Class	Variable	Label	Description	Format Text
11. Process	**mortality_review_dt_adj**	Adjusted Date of Mortality Review?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.F="No Form" .M="Missing Value" 0="No"

Section 8: Cancer Treatment

Class	Variable	Label	Description	Format Text
01. Principal	**cncr_surgery**	Surgery	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 1="Single Lobectomy" 2="Bilobectomy" 3="Simple Pneumonectomy" 4="Extended Lobectomy" 5="Thoracotomy and Biopsy" 6="Wedge Resection"
01. Principal	**primary_trt**	Primary Treatment	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 0="None" 1="Complete Resection" 2="Incomplete Resection" 3="Radiation Only" 4="Chemotherapy Only" 5="Radiation and Chemotherapy" 6="Cautery"
01. Principal	**supplementary_trt**	Supplementary Treatment	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 0="None" 3="Radiation Only" 4="Chemotherapy Only" 5="Radiation and Chemotherapy"
01. Principal	**treatment_days**	Days from Randomization to Date of Treatment	Populated only for participants with probable or confirmed lung cancer.	Numeric .F="No Form"
01. Principal	**treatment_loc**	Location of Treatment	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" 1="Memorial Hospital" 2="Other"
11. Process	**treatment_dt_adj**	Adjusted Date of Treatment?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.F="No Form" 0="No"

Section 9: Mortality Status

Class	Variable	Label	Description	Format Text
01. Principal	**death_days**	Days from Randomization to Date of Death	The date on the follow-up questionnaire for deceased participants.	Numeric .F="No Form" .M="Missing Value"
01. Principal	**death_flag**	Death Flag	Does the trial have any indication that the participant is dead (if the only indication is from the mortality review, which is conducted on all lung cancers, then they are considered alive).	0="Alive" 1="Lung Cancer Death" 2="Other Cause"
01. Principal	**death_icda_best**	Deceased: Best ICDA Code		Char, 5
01. Principal	**death_icda_underlying**	Deceased: Underlying ICDA Code		Char, 5
01. Principal	**death_panel_cat**	Deceased: Death Panel Category		.F="No Form" .M="Missing Value" 12="Lung Cancer Not Confirmed - Death Related To Study Initiated Procedures" 13="Lung Cancer Never Confirmed - Death From Other Causes" 21="First Lung Cancer Was Cured; Died From a 2nd Lung Cancer" 23="Lung Cancer No Longer Present - Death From Other Causes" 31="Lung Cancer Present - Death From Primary Lung Cancer" 32="Lung Cancer Present - Death Related To Study Initiated Procedures" 33="Lung Cancer Present - Death From Other Causes" 41="Lung Cancer Probable - Death From Primary Lung Cancer" 43="Lung Cancer Probable - Death From Other Causes" 53="Cancer In Thorax (Not Primary) - Death From Other Causes"
01. Principal	**mortality_exitdays**	Days from Randomization to Date of Death or Last Follow-Up	The date of death or last follow-up is the date of death if present, otherwise it is the latest date from the follow-up status record, follow-up non status record, cytology record, radiology record, or date of diagnosis, or date of last contact.	Numeric .M="Missing Value"

Class	Variable	Label	Description	Format Text
01. Principal	**mortality_review_cat**	Death Category as in Mortality Review	Populated only for participants with probable or confirmed lung cancer.	.F="No Form" .M="Missing Value" 21="First Lung Cancer Was Cured; Died From a 2nd Lung Cancer" 23="Lung Cancer No Longer Present - Death From Other Causes" 31="Lung Cancer Present - Death From Primary Lung Cancer" 32="Lung Cancer Present - Death Related To Study Initiated Procedures" 33="Lung Cancer Present - Death From Other Causes" 41="Lung Cancer Probable - Death From Primary Lung Cancer" 43="Lung Cancer Probable - Death From Other Causes" 99="Unknown Code (BC)"
09. Secondary	**death_panel_cat_flag**	Deceased: Death Panel Category Character Associated with Extra Asterisk Character?	Does the original death panel character variable have an asterisk character at the end?	0="No" 1="Yes"
11. Process	**death_dt_adj**	Adjusted Date of Death?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.F="No Form" .M="Missing Value" 0="No" 1="Yes"
11. Process	**mortality_exitdt_adj**	Adjusted Date of Death or Last Follow-Up?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.M="Missing Value" 0="No" 1="Yes"

Section 10: Follow-Up Questionnaire

Class	Variable	Label	Description	Format Text
01. Principal	qx_chest_xray1-27	Follow-Up Questionnaire: Chest X-Ray		.F="No Form" .N="Not Applicable" 0="No Response" 1="Chest X-Ray at Pmi-Strang Clinic" 2="Chest X-Ray at Hospital" 3="Chest X-Ray at Doctor's Office" 4="No Chest X-Ray this Year" 9="Invalid Response"
01. Principal	qx_counter	Number of Entries for Follow-Up Questionnaire		Numeric .F="No Form"
01. Principal	qx_days1-27	Days from Randomization to Date of Follow-Up Questionnaire		Numeric .F="No Form" .M="Missing Value" .N="Not Applicable"
01. Principal	qx_hospitalized1-27	Follow-Up Questionnaire: Hospitalized		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
01. Principal	qx_seen_physician1-27	Follow-Up Questionnaire: Seen A Physician		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
01. Principal	qx_stopped_smoke1-27	Follow-Up Questionnaire: Stopped Smoking		.F="No Form" .N="Not Applicable" 0="No" 1="Yes"
11. Process	qx_dt_adj1-27	Adjusted Date of Follow-Up Questionnaire?	Indicates whether the respective date was adjusted. If a month does not exist within [1:12], it is adjusted to a potentially valid month if it is a multiple of 10; otherwise to July. If the day does not exist within [1:31], it is adjusted to the correct maximum of days if it is near the correct maximum; otherwise, to the 15th.	.F="No Form" .M="Missing Value" .N="Not Applicable" 0="No" 1="Yes"