

Supplement 2

Pan-European Lung Cancer Dataset

Basic patient features		
Data item	Definition	Detailed definition
Date of birth	dd/mm/yyyy	
Sex	Male or female	1=male 2=female 999=undisclosed/unknown
Country of registration	ISO-3166	2 letter code
Education level	Numerical value (ICHOM)	Indicate highest level of education completed 0=None 1=Primary 2=Secondary 3=Tertiary (college, university) 999=Don't know
Date of referral	dd/mm/yyyy	Date on which referral made with respect to potential lung cancer. This could include, self-referral, primary to secondary care, within secondary care Option for missing/unknown.
Date of diagnosis	dd/mm/yyyy	Date the first histopathology/cytology sample was taken which confirmed malignancy, If date of histopathology sample not available then index date based on following criteria in descending order. (as per IARC) 1. Date of first histological or cytological confirmation of this malignancy (with the exception of histology or cytology at autopsy). This date should be, in the following order: a) date when the specimen was taken (biopsy) b) date of receipt by pathologist c) date of the pathology report 2. Date of admission to hospital because of this malignancy. 3. When evaluated at an outpatient clinic only: date of first consultation at the out-patient clinic because of this malignancy. 4. Date of diagnosis other than 1,2 or 3. 5. Date of death, if no information is available other than the fact that the patient died of a malignancy. 6. Date of death, if the malignancy is discovered at autopsy.
Date of final pathology report	dd/mm/yyyy	Date of final pathology report to include molecular analysis where appropriate Option for missing/unknown.
Mode of presentation	How was lung cancer first suspected? Numerical value	0=screening 1=symptoms 2=incidental finding 3=other (free text box to specify) 999=don't know

Basis of diagnosis	Numerical value (ICHOM)	1=Clinical 2=Histology 3=Cytology 999=unknown
	Clinical	Diagnosis made before death with or without diagnostic techniques (e.g. X-rays, endoscopy, imaging, ultrasound, exploratory surgery) but without a tissue diagnosis
	Histology	Histological examination of tissue from the primary tumour or metastasis, including all cutting and bone marrow biopsies. Also includes autopsy specimens.
	Cytology	Examination of cells whether from a primary or secondary site, including fluids aspirated using endoscopes or needles. Also including microscopic examination of peripheral blood films and trephine bone marrow aspirates
Tumour details		
Data item	Definition	Detailed definition
Histology	System used	International Classification Diseases -Oncology-3 rd edition (ICD-O-3 rd edition) (covers the entire 2015 World Health Organization (WHO) Classification of Tumours of the Lung) Systematized Nomenclature of Medicine (SNOMED) Based on which system is used, a list of possible options will appear, and the correct histology field can be ticked.
	If no recognised system used then drop down menu appears with a limited list	
	1	Small cell carcinoma
	2	Non-small cell carcinoma (NSCLC) NOS
	3	Squamous cell carcinoma
	4	Adenocarcinoma
	5	Large cell neuroendocrine carcinoma
	6	Carcinoid-typical
	7	Carcinoid-atypical
	8	Adenocarcinoma in situ
	9	Spindle/pleiomorphic/giant cell NSCLC
	10	Other (free text box appears)
	999	Unknown
Molecular analysis*	Was this performed?	0=No 1=Yes 999=don't know
Further questions only relevant if molecular analysis performed.	EGFR mutation (ICHOM) Numerical value	Indicate presence of EGFR activating mutation 0=No 1=Yes 2=Failed analysis 999=don't know
	If activating mutation found, on which exon?	0=not relevant 1=exon 18 2=exon 19 3=exon 21 999=don't know
	EGFR mutation T790M	Indicate presence of EGFR mutation of resistance 0=No

		1=Yes 2=Failed analysis 999=don't know
	ALK translocation (ICHOM) Numerical value	Indicate presence of ALK translocation 0=No 1=Yes 2=failed analysis 999=don't know
	Ros 1	Indicate presence of Ros1 translocation 0=No 1=Yes 2=failed analysis 999=don't know
	BRAF	Indicate presence of BRAF mutation 0=No 1=Yes 2=failed analysis 999=don't know
	PD-L1 status	Indicate PD-L1 status 0=Not expressed 1=Some expression 2=failed analysis 999=don't know
	PD-L1 percent expression	0=not applicable (ie 0 above) 1=<1 2=1-9.9% 3=10-49% 4=>50%
	RET	Indicate presence of RET translocation 0=No 1=Yes 2=failed analysis 999=don't know
	MET	Indicate presence of MET amplification 0=No 1=Yes 2=failed analysis 999=don't know
	MET exon 14	Indicate presence of MET mutation exon 14 0=No 1=Yes 2=failed analysis 999=don't know
	HER 2	Indicate presence of HER2 mutation 0=No

		1=Yes 2=failed analysis 999=don't know
* Annual updates expected as molecular medicine develops		
Stage	Final pre-treatment clinical stage	
Basis of stage	What method was used to decide the final pre-treatment clinical stage? Numerical value	0=imaging only 1=imaging AND non-surgical pathology samples 2=imaging and surgical biopsies (mediastinoscopy, VATS procedure) 999=Don't know <u>Definition of non-surgical samples;</u> EBUS, EUS, percutaneous lung or pleural biopsy, pleural aspiration, bronchoscopy.
Investigations performed?	CT scan Numerical value	0=No 1=Yes 999=Don't know
	PET-CT	0=No 1=Yes 999=Don't know
	Bronchoscopy	0=No 1=Yes 999=Don't know
	EBUS	0=No 1=Yes 999=Don't know
	EUS	0=No 1=Yes 999=Don't know
	Mediastinoscopy	0=No 1=Yes 999=Don't know
	Histopathology or cytology from node outside chest	0=No 1=Yes 999=Don't know
	Sampling of pleura or pleural fluid aspiration (medical)	0=No 1=Yes 999=Don't know
	VATS thoracoscopy	0=No 1=Yes 999=Don't know
	Imaging of metastasis (e.g.; CT/MRI brain, MRI spine, MRI adrenal etc)	0=No 1=Yes 999=Don't know
	Histopathology of metastasis (eg liver biopsy)	0=No 1=Yes 999=Don't know
	Exploratory open thoracic	0=No

	surgery	1=Yes 999=Don't know
Tumour size	Numerical value	The longest single direction size (cm to one decimal point, e.g.3.2)
Staging system	Which staging system has been used?	UICC 7th or 8th Based on this answer; drop down menu appears for T, N and M stage.
Tumour stage	Mixed value	UICC Version 7; 1a through to 4 UICC Version 8; 1mi through to 4 999=Unknown/X
Nodal stage	Mixed value	Version 7 or 8; 0 to 3 999=unknown
Extent of N1 disease	Numerical value	0=Not applicable 1=Single station N1 disease 2=Multi-station N1 disease
Extent of N2 disease	Numerical value	0=Not applicable 1= microscopic N2 node found at final pathological (post-operative) specimen, 2= Single station N2 node without N1 disease ('skip' lesion) 3= Single station N2 node with N1 involvement 4= Multi-station N2 disease 5= Bulky or fixed multi-station N2 disease
Metastasis stage	Mixed value	Version 7: 0 through to 1b Version 8; 0 through to 1c 999=unknown/X
Number of metastatic lesions	Numerical value	0=no metastatic spread (ie M0 above) 1, 2, 3 onwards 999=don't know
Site of metastases	Liver Numerical value	0=no 1=Yes If yes, specify number of metastatic lesions 1, 2, 3 onwards 999=don't know
	Brain	0=no 1=Yes If yes, number of metastatic lesions 1, 2, 3 onwards 999=don't know
	Adrenal	0=no 1=Yes If yes, specify number of metastatic lesions 1, 2, 3 onwards 999=don't know
	Bone	0=no 1=Yes If yes, specify number of metastatic lesions 1, 2, 3 onwards 999=don't know
	Other	Free text box to confirm site of spread

Extended patient features		
Data item	Definition	Detailed definition
Performance status (final pre-treatment)	ECOG (WHO)	Numerical value 0-4 999=unknown
	0	Able to carry out all normal activity without restriction
	1	Restricted in physically strenuous activity, but able to walk and do light work
	2	Able to walk and capable of all self-care, but unable to carry out any work. Up and about more than 50% of waking hours
	3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
	4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair
	999	Unknown/not recorded
Smoking status	Numerical code (ICHOM)	1=never smoker (<100 cigarettes ever) 2=ex-smoker (stopped at least 1 year before inclusion, ie diagnosis) 3=current smoker 999=don't know
Smoking pack years	Numerical value	Simply the number of pack years smoked, regardless of ex or current smoker, Eg 20, 40, etc 999=don't know
Comorbidity at baseline From medical consultation with patient	ICHOM modified Self-administered Comorbidity Questionnaire (SCQ; Sangha et al 2003) Drop down menu; multiple options possible	Have you been told by a doctor that you have any of the following: 0=I have no other diseases 1=Heart disease (eg, angina, heart attack or heart failure) 2=High blood pressure 3=Leg pain when walking due to poor circulation 4=Lung disease (eg, asthma, chronic bronchitis, COPD, or emphysema) 5=Diabetes 6=Kidney disease 7=Liver disease 8=Problems caused by stroke 9=Disease of the nervous system (eg, Parkinson disease, multiple sclerosis) 10=other cancer (within the last 5 years) 11=Depression 12=Arthritis
Weight	Numerical value	In kg. 999=don't know
Height	Numerical value	In m 999=don't know
Lung function (at baseline)	Numerical value (ICHOM)	Observed FEV ₁ (L) (e.g. 1.35) 999=don't know
	Numerical value (ICHOM)	Percent predicted FEV ₁ (e.g. 56, would represent 56% predicted) 999=don't know
	Numerical value	Observed FVC (L) (e.g.. 2.3) 999=don't know
	Numerical value	Percent predicted transfer factor (KCO) (e.g. 85 would represent 85% predicted)

		Numerical value 999=don't know
Quality of Life		
At diagnosis	EORTC QLQ-C30	Score (maximum value 126)
At end of 1 st line treatment	EORTC QLQ-C30	Score (maximum value 126)
At 6 months post diagnosis	EORTC QLQ-C30	Score (maximum value 126)
Lung cancer outcomes		
Data item	Definition	Detailed definition
Lung Cancer Nurse Specialist (LCNS)	Contact made with LCNS	0=no 1=Yes 999=Don't know
Treatment		
Treatment intent (reflects the intent of the treating physician or MDT)	Numerical code	1=curative intent 2=non-curative intent 3=no active treatment 999=don't know
	Curative	This is single or multimodality treatment which is hoped will remove the threat of lung cancer on the patient's life expectancy.
	Non-curative	This is single or multimodality treatment which is expected to gain local control, or limit the progression of the disease, but unlikely to remove the threat of lung cancer on the patient's life expectancy.
	No active treatment	This would be those patients who decline, or are too frail for, radiotherapy or chemotherapy and simply receive medication for symptom control or a watch and wait policy
First line treatment given to primary tumour		
	Numerical code	Choose one option only from list below.
	1	Surgery alone
	2	Hyperfractionated radiotherapy
	3	External beam radiotherapy (curative intent but not CHART)
	4	Stereotactic Radiotherapy
	5	Radiofrequency/microwave ablation
	6	Brachytherapy
	7	Palliative radiotherapy to lung primary
	8	Concurrent chemo-radiotherapy
	9	Sequential chemo-radiotherapy
	10	Induction radiotherapy (pre surgery)
	11	Induction chemotherapy (pre surgery)
	12	Palliative Chemotherapy
	13	Targeted/biological therapy (TKI etc)
	14	Immunotherapy
	15	Interventional bronchoscopy
	16	Specialist palliative care
	17	Other (free text)
	999	Don't know/not recorded
First line treatment not given or change in treatment plan		0=Not relevant

		1=Patient declined 1 st line treatment offered 2=Patient deteriorated and no longer eligible for 1 st line treatment 3=Hospital unable to provide 1 st line treatment 4= other (free text)
Date of 1st line treatment	dd/mm/yyyy	Date of the start of 1 st line treatment, ie date of operation, or first day of radiotherapy or chemotherapy regime, or appointment with specialist palliative care physician.
Type of operation Additional question only if option 1 above	Numerical value	0=incomplete resection (residual macroscopic disease evident) 1=segmentectomy 2=wedge resection 3=lobectomy 4=bi-lobectomy 5=sleeve lobectomy 6=pneumonectomy 999=don't know
Pathological stage Additional question only if option 1 above	Staging system used	IASLC version 7 or 8
	Pathological stage	Based on version used can then have drop down menu for pathological stage
Nature of radiotherapy Additional questions only if options 2,3,4,7 or 10 chosen	Total dose given (Grey)	Absolute number.
	Number of fractions	Absolute number (eg 6)
	Number of days or radiotherapy treatment	Absolute number (eg. 12)
Nature of chemotherapy Additional question only if options 11 or 12 chosen	Numerical value	1=single agent chemotherapy 2= Doublet platinum-based chemotherapy 3=other (free text) 999=don't know
Additional supportive 1st line treatment given?		
	Numerical value	0=No 1=Yes If yes then further question appears
Type of additional supportive 1st line treatment		
	Numerical value	
	1	Stereotactic radiotherapy to brain metastases
	2	Radiotherapy for spinal cord compression
	3	Prophylactic Cranial Irradiation
	4	Whole brain radiotherapy
	5	Radiotherapy for oligometastases
	6	SABR for oligometastases
	7	Radiotherapy for SVCO
	8	Radiotherapy to mediastinum
	9	Specialist palliative care
	10	Surgical resection of metastases

	11	Pleural intervention (see below)
	12	Other (free text)
Date of first radiotherapy session	dd/mm/yyyy	
Nature of radiotherapy (1-8)	Total dose given (Grey)	Absolute number (e.g. 30)
	Number of fractions	Absolute number (e.g. 6)
	Number of days or radiotherapy treatment	Absolute number (e.g. 12)
Date of palliative care	dd/mm/yyyy	Date of first appointment with specialist palliative care physician
Date of surgery	dd/mm/yyyy	
Date of pleural intervention	dd/mm/yyyy	
Type of Pleural intervention	Numerical value	1=thoracocentesis 2=chest drain 3=pleurodesis 4=indwelling chest drain
How is the patient followed up after 1 st line treatment?	Numerical value, please pick single item from list. Options are ranked in descending order. If multiple answers apply, please pick the first answer in the list	1=Regular out-patient visits with physician (member of MDT) 2=Follow up with lung cancer nurse specialist 3=Virtual follow-up after imaging 4=Phone contact with patient 5=Referred back to primary care doctor 0=No follow-up 999=don't know
Date of completion of 1 st line treatment	dd/mm/yyyy	
Response to 1 st line treatment?	Numerical value	0=Complete remission 1=Partial response 2=Stable disease 3=Progression 999=don't know
Date of relapse	dd/mm/yyyy	
How was relapse detected?		0=Planned imaging 1=Symptoms 2=Incidental finding with unrelated problem 999=don't know
Subsequent treatment to lung primary		
	Numerical code	More than one treatment option can be chosen during the patient treatment programme (please confirm with dates below)
	1	Surgery
	2	Chemotherapy and radiotherapy in addition to surgery (tri-modality treatment)
	3	Hyperfractionated radiotherapy
	4	External beam radiotherapy (curative intent but not CHART)
	5	Stereotactic Radiotherapy (3-8 fractions)
	6	Radiofrequency/microwave ablation
	7	Brachytherapy
	8	Palliative radiotherapy to lung primary
	9	Concurrent chemo-radiotherapy

	10	Sequential chemo-radiotherapy
	11	Palliative Chemotherapy
	12	Targeted/biological therapy (TKI etc)
	13	Immunotherapy
	14	Interventional bronchoscopy
	15	Specialist palliative care
	999	Don't know/not recorded
Date of surgery	dd/mm/yyyy	
Date of first radiotherapy session	dd/mm/yyyy	
Nature of radiotherapy	Total dose given (Grey)	Absolute number (e.g. 30)
	Number of fractions given	Absolute number (e.g. 6)
	Number of days or radiotherapy treatment	Absolute number (e.g. 12)
Date first chemotherapy started	dd/mm/yyyy	
Date of last chemotherapy dose	dd/mm/yyyy	
Date of interventional bronchoscopy	dd/mm/yyyy	
Date of specialist palliative care	dd/mm/yyyy	Date of first appointment with specialist palliative care physician
Clinical trial	Is the patient part of a clinical trial?	0=No 1=Yes 999=Don't know
Date of death	dd/mm/yyyy	

Legend: data items in black are mandatory within the proposed minimum dataset; those in blue are desirable dataset.

Dd/mm/yyyy; Date, month and year.