PRIVILEGED COMMUNICATION

The International Association for the Study of Lung Cancer (IASLC)
Lung Cancer Staging Project

PROTOCOL
FOR PURPOSE OF GRANT APPLICATION AND ETHICS REVIEW

IASLC Staging and Prognostic Factors Committee, Full Committee Chair
Hisao Asamura, MD
Keio University School of Medicine, Tokyo, Japan

IASLC Staging and Prognostic Factors Committee, Full Committee Vice Chair
Kemp Kernstine, MD, PhD
The University of Texas Southwestern Medical Center, Dallas, Texas, United States of America

T Subcommittee Chair
Ayten K Cangir, MD
Ankara University School of Medicine, Ankara, Turkey

N Subcommittee Chair
James Huang, MD
Memorial Sloan-Kettering Cancer Center, New York, New York, United States of America

M Subcommittee Chair
Kwun Fong, MBBS, FRACP, PhD
Prince Charles Hospital, Brisbane, Australia

Ground Glass Opacities (GGO) and Adenocarcinoma In Situ (AIS) Subcommittee Chair
William Travis, MD
Memorial Sloan-Kettering Cancer Center, New York, New York, United States of America

Neuroendocrine Tumors Subcommittee Chair
Ming Tsao, FRCPC, MD
Princess Margaret Cancer Centre, Toronto, Ontario, Canada

Stage Subcommittee Chair
Hisao Asamura, MD
Keio University School of Medicine, Tokyo, Japan
**Lymph Node Chart Subcommittee Chair**
Shun-ichi Watanabe, MD
National Cancer Center Hospital, Tokyo, Japan

**Validation and Methodology Subcommittee Chair**
Frank Detterbeck, MD
Yale University School of Medicine, New Haven, Connecticut, United States of America

**Prognostic Factors Subcommittee Chair**
Frank Detterbeck, MD
Yale University School of Medicine, New Haven, Connecticut, United States of America

**R Factor Subcommittee Chair**
John Edwards, MD
University of Leicester, Glenfield Hospital, Leicester, United Kingdom

**Molecular Database Taskforce Chair**
David Carbone, MD, PhD
Ohio State University, Columbus, Ohio, United States of America

**Statistics and Data Management**
John Crowley, PhD
Cancer Research and Biostatistics (CRAB), Seattle, Washington, United States of America
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROCOTOL ......................................................................................................................... 1</td>
</tr>
<tr>
<td>TABLE OF CONTENTS ..................................................................................................... 3</td>
</tr>
<tr>
<td>1.0 OBJECTIVES .............................................................................................................. 4</td>
</tr>
<tr>
<td>2.0 BACKGROUND ............................................................................................................. 5</td>
</tr>
<tr>
<td>3.0 STUDY DESIGN ......................................................................................................... 6</td>
</tr>
<tr>
<td>4.0 SAMPLE SELECTION ............................................................................................... 7</td>
</tr>
<tr>
<td>5.0 INTERVENTION ........................................................................................................ 7</td>
</tr>
<tr>
<td>6.0 ELIGIBILITY CRITERIA ............................................................................................. 8</td>
</tr>
<tr>
<td>7.0 DESCRIPTIVE FACTORS .......................................................................................... 8</td>
</tr>
<tr>
<td>8.0 STATISTICAL CONSIDERATIONS ............................................................................... 8</td>
</tr>
<tr>
<td>9.0 DE-IDENTIFICATION OF DATA .................................................................................. 9</td>
</tr>
<tr>
<td>10.0 SITE APPLICATION MATERIALS ............................................................................... 10</td>
</tr>
<tr>
<td>11.0 PROPERTY OF THE DATA BASE AND PUBLICATION POLICY .............................. 10</td>
</tr>
<tr>
<td>12.0 DATA COLLECTION PROCESSES .......................................................................... 10</td>
</tr>
<tr>
<td>13.0 OVERSIGHT BY VALIDATION AND METHODOLOGY COMMITTEE ................. 11</td>
</tr>
<tr>
<td>14.0 SECONDARY USE OF THE IASLC LUNG CANCER DATA BASE .......................... 11</td>
</tr>
<tr>
<td>15.0 REFERENCES .......................................................................................................... 13</td>
</tr>
</tbody>
</table>
1.0 OBJECTIVES

1.1. The primary objectives of this study are:
   1.1.1. To assess the prognostic value and validity of each component of the eighth edition of the tumor, node, metastasis (TNM) classification for lung cancer with respect to the overall survival of patients with newly diagnosed lung cancer.
   1.1.2. To identify and validate additional descriptors for possible inclusion in future revisions to the TNM classification.

1.2. T-Component objectives:
   1.2.1. To assess the prognostic impact of tumor size
   1.2.2. To assess the classification capacity of each descriptor defining T-status.
   1.2.3. To study new conditions not included in the present T (e.g., differences between parietal pleura invasion and rib invasion).

1.3. N-Component objectives:
   1.3.1 To assess the prognostic impact of N-status.
   1.3.2 To assess the prognostic impact of:
      a. Nodal extent (single vs multiple station involvement in N1 and N2 locations),
      b. Nodal size, i.e. the largest involved node within the relevant N category, and
      c. Individual nodes being involved in each nodal category.
   1.3.3 To assess the prognostic impact of extracapsular extension.
   1.3.4 To assess the prognostic impact of the N3 nodal location, i.e. contralateral mediastinum, ipsilateral or contralateral supraclavicular fossa.

1.4 M-Component objectives:
   1.4.1 To assess the prognostic impact of M-status.
   1.4.2 To assess the prognostic impact of:
      1.4.2.1 Single metastasis in a single organ
      1.4.2.2 Multiple metastases in a single organ, and
      1.4.2.3 Multiple metastases in several organs.

1.5 Objectives regarding other prognostic factors:
   1.5.1 To assess the prognostic impact of histologic type and grade.
   1.5.2 To assess the reliability of staging methods utilized in clinical staging (for those tumors with pre-treatment and post-surgical classification).
   1.5.3 To assess the prognostic impact of complete, incomplete, and uncertain resections, according to the proposed definitions of the IASLC.
   1.5.4 To assess the prognostic impact of clinical factors, including co-morbidity and pulmonary function tests.
   1.5.5 To assess the prognostic impact of maximum standard uptake value (SUV max), at the primary site and in any positive nodal sites, for those patients with positron emission tomography (PET) scans in the pre-treatment staging.
2.0 BACKGROUND

The objectives of the first iteration of the Lung Cancer Staging Project of the International Association for the Study of Lung Cancer (IASLC)\(^1\) were achieved in 2007 with the submission of recommendations for the seventh edition of the tumor, node, metastasis (TNM) classification for lung cancer to the International Union Against Cancer (UICC) and to the American Joint Committee on Cancer (AJCC). The UICC and AJCC accepted these recommendations. These core recommendations and the methodology used in the analysis of the retrospective database were published\(^2\), \(^3\), \(^4\), \(^5\), \(^6\) in 2007, along with additional publications on small cell lung cancer, carcinoid tumors, and prognostic factors\(^7\), \(^8\), \(^9\).

The limitations of the analysis of the retrospective database derived from the fact that most databases that contributed cases to the international database were not designed to study the TNM classification of lung cancer. The most important consequence was that, while the clinical or pathologic T status was recorded in most of the databases, few included the finer details, such as the specific anatomic sites of tumor extension. For this reason, most of the descriptors that define T3 and T4 tumors could not be validated in this retrospective study\(^2\). The same was true for the potential subdivision of the N1 and N2 nodal spread based on the number of involved nodes/nodal stations or nodal zones\(^3\), and for the differences in the various forms of M1 disease\(^4\). In addition, subtle differences between nodal maps used in different parts of the world – e.g. the Mountain and Dressler 1997 modification to the American Thoracic Society (ATS)\(^10\) map and the Naruke-Japan Lung Cancer Society map\(^11\), \(^12\) – complicated previous attempts to analyze international data on nodal involvement.

To overcome the limitations of the previous, retrospectively amalgamated database, the Staging and Prognostic Factors Committee (SPFC) of the IASLC launched a web-based Electronic Data Capture (EDC) system for staging and survival data in 2009\(^13\) with the general objective to refine future editions of the TNM classification for lung cancer. The EDC was designed to facilitate validation of all T, N, and M descriptors, with special attention to those that could not be validated with the analysis of the retrospective database, and the investigation of the prognostic value of other descriptors of interest to the SPFC that were not included in the TNM classification.

As a result of this web-based data collection effort, supplemented by other large, external data sources, a database of 94,708 patients diagnosed around the world from 1999-2010 was constructed, from which the IASLC SPFC developed recommendations toward the eighth edition of the TNM staging system. These recommendations were published in 2015\(^14\), \(^15\), \(^16\), \(^17\), \(^18\), \(^19\) and were accepted by the UICC and AJCC in 2017 (with AJCC implementation effective in 2018).

As the project enters its third cycle, with the goal of developing recommendations for the 9\(^{th}\) edition of TNM, its continued success will depend on the extent of the
international participation and the quality of the data. The initial retrospective staging project showed that quality of the data is even more important than its size. In the second iteration of the project, the superior quality of TNM descriptor data from the EDC relative to retrospectively amalgamated data was clearly demonstrated. The subset of cases that were entered via the EDC was instrumental in the development of the 8th edition recommendations due to the level of detail and consistency of the submitted data. For example, the distinction between cases with multiple distant metastatic lesions and those with a single distant metastasis could not have been identified without the EDC dataset.

Useful information not related to the anatomic extent of the disease can also be derived from the newly revised data elements. Including the methods used in clinical staging allows exploration of their reliability in those patients undergoing lung resection, in whom the pre-treatment and post-surgical classifications can be compared. The IASLC has published consensus guidelines on clinical staging based on the best evidence available in clinical practice21, 22, 23. With continued collection of the data elements related to clinical staging methods, the previously developed guidelines can be validated, and future staging recommendations can be made in the context of adherence to those guidelines.

One of the objectives of the TNM classification is to assign a prognosis based on the anatomical extent of the disease. However, there are other factors that influence prognosis of lung cancer that are not related to its anatomic extension. Sex, age and comorbidity24, 25, 26, biological parameters27, and molecular and genetic factors are known to influence prognosis, but have never been integrated, along with the TNM classification, into a valid, clinically useful prognostic system. Information on comorbidity and basic blood analyses is easily available from most patients. The maximum standard uptake value, which has shown prognostic relevance28, will also be registered in those patients undergoing PET scan in the pre-treatment staging of their tumors. The importance of molecular and genetic factors is now undisputed, both in terms of survival prognosis as well as interactions with treatment. Recognizing this importance, the collection of detailed biomarker information and specific systemic treatments, with emphasis on targeted agents, is a new feature of the revised EDC. This data collection system is designed to be easily expanded as new biomarkers and discovered, and as new drugs come to market.

Centralized collection of all these anatomic and non-anatomic parameters has been found to be most effective in addressing the research questions of the SPFC of the IASLC29. This document is provided to collaborating institutions so that we may standardize the processes and procedures for conducting this study across multiple institutions.

3.0 STUDY DESIGN

This is an international, multi-institutional cohort study that will collect detailed information on the extent of disease, personal and demographic characteristics, comorbid illness status, treatment and survival of newly-diagnosed lung cancer patients.
Ideally, an inception cohort will be enrolled prospectively at each site, and data will be collected using a standardized abstraction tool. However, because it is unlikely that accrual goals will be met using this option alone, sites may alternatively petition the SPFC to transfer data from an existing database.

Data completeness and logic checks will be conducted on an ongoing basis. Analyses will be conducted at CRAB. Each participating institution will have access to their own patients’ data and will be eligible to conduct secondary investigations of the larger database subject to approval by the IASLC.

4.0 SAMPLE SELECTION

Based on the experience from the original retrospective study, three types of study samples will be targeted, depending on the nature of the collaborating institution: population-based, institution-based and clinical series. In each case, the intention is to describe the experience of an unselected group of patients. Sample selection for each is described below.

4.1 Population-based sample selection will likely involve enhancement of a population-based cancer registry with the data elements required for this study. All patients diagnosed within the study period may be included, or a random sample from the registry within the study period may be included. Documentation of the population coverage of the registry will be required for a sample to fall into this category.

4.2 Institution-based sample selection will likely involve the capture of information on all newly-diagnosed lung cancer patients seen at that institution during the period of the study. Usually, this involves the use of an institution’s tumor registry that will be enhanced with the data elements required for this study. Description of the institution’s referral pattern will be required.

4.3 Clinical series sample selection will capture information on an inception cohort of all newly-diagnosed patients presenting to a defined clinical service during the period of the study. All such patients will be tracked with documentation regarding data completeness and losses to follow up.

In considering applications for participation in the project, the SPFC will grant preference to sites which implement one of the above methods of sample selection.

5.0 INTERVENTION

Subjects will not be assigned to any specific intervention as a result of inclusion in this observational data base.
6.0 ELIGIBILITY CRITERIA

6.1 Subjects must have newly diagnosed non-small or small cell bronchogenic carcinomas, including neuroendocrine and carcinoid tumors of the lung.

6.2 Lung cancer must be confirmed by histology or cytology, with a diagnosis date no earlier than January 1, 2011.

6.3 For a subject to be eligible for inclusion, there must be sufficient information available to classify the subject according to the eighth edition of the TNM classification for lung cancer.

7.0 DESCRIPTIVE FACTORS

Patients will be described by pretreatment T, N, and M status, treatment (surgically managed vs not), by country of origin and study sample type. Enrollment will be monitored with respect to these descriptive factors, with two objectives in mind: 1) to track recruitment of specific subgroups defined by geography, stage, or treatment modality with a view to targeting additional institutions and/or clinical settings if under-representation exists and 2) to demonstrate that the study sample is unbiased with regard to subject selection.

8.0 STATISTICAL CONSIDERATIONS

Participation in the previous two database cycles was high, with approximately 100,000 cases submitted for each of the last two revisions. During the most recent effort, 4,631 cases were submitted via electronic data capture. Based upon the number of new participants expressing interest and the expectation of past participants to continue using the EDC or to convert to it, we anticipate a larger proportion of cases to be submitted in this fashion for the next revision. The data derived from EDC submission are uniformly complete, and therefore the calculations below consider only the number of cases that we expect to accrue via the EDC.

The EDC accrual expectation is 20,000 non-small cell lung cancer (NSCLC) cases diagnosed between the beginning of 2011 and the end of 2020, and a smaller proportion of small cell lung cancer (SCLC) and other histologic types. This database is not a population-based registry and the stage distribution is not expected to reflect the distribution of lung cancer stage in the general population. Based on the previous database, we expect the stage distribution for NSCLC to be approximately 40% stage I, 10% stage II, 25% stage III, and 25% stage IV. For each of the existing ten stage groups of the 8th edition (IA, IB, IC, IIA, IIB, IIA, IIIB, IIIC, IVA, IVB), the smallest anticipated stage groups would be the stage IIA and IIB, with approximately 1000 cases in each of these two group from the EDC.

Previously, the 24 month overall survival rate in stage IIA was 81%, and 75% in the stage IIB (HR=1.2) according to the full NSCLC analysis set in the IASLC database. (The
difference was similar in the National Cancer Database, although the absolute survival rates were lower.) A difference smaller than this would not be expected to warrant separate stage categories. In a comparison of two groups of 1000 cases each, after 10 years of accrual and an additional 2 years of follow-up, and assuming exponentially distributed survival times, there would be 92% power to detect a hazard ratio of 1.2 with an alpha level of 5% in a one-sided test. This is a worst case scenario as we expect the stage II to be the smallest group of subjects. For example in the stage III, we expect 1600 cases in each of 3 groups. If we wish to detect at minimum a hazard ratio of 1.2 between the potential stage IIIA and IIIB, we would have 99% power to do so at an alpha level of 5% in a one-sided test. Applying a Bonferroni correction to account for multiple comparisons under the assumption of 9 separate between-group tests, power would range from 70% to 92% for the above scenarios.

These power calculations address the simplified case of formally confirming up to 10 overall stage categories. Initially, more exploratory analyses will be conducted to inform the stage categories, and subsequently other subsets will be separately considered, such as the SCLC and neuroendocrine histologic types. Additionally, we anticipate a substantial number of cases submitted by participating sites will be transferred to the project rather than entered directly into the EDC system. These datasets will vary in terms of data elements, and some may only be used to answer some of the questions that arise as part of the initial analyses. Although some of these datasets will be sufficiently complete and utilized in the primary analyses, we do not include them in the power calculations.

9.0 DE-IDENTIFICATION OF DATA

All sites must agree to gather identifiable private information of research subjects in compliance with applicable law and with respect and regards for human subjects. Each participating institution will secure approval of the project from their local Research Ethics Board.

Participating sites must agree unequivocally to prohibit release of individually identifiable private data to CRAB for research purposes. CRAB will receive only ‘coded’ data for analysis. The ‘coded’ data sent to CRAB must not be able to be linked to individual research subjects, either directly or indirectly through the coding system, by any member of CRAB’s research team. Where personal identifiers might inadvertently be included with data received, CRAB will delete/destroy this identified data, and immediately notify the site to replace with de-identified data.

If ever visiting the site, CRAB staff may access or utilize individually private information but these activities become subject to the oversight of the site’s Institutional Review Board. At no time will CRAB employees record any private information.

CRAB, as an institution, is not considered to be “engaged” in human subjects research for this project.
10.0 SITE APPLICATION MATERIALS

This section includes the necessary application materials for any site interested in contributing data to the staging project. These application materials and other supportive documentation for the project can be accessed online at https://iaslc.crab.org/LC/LCStagingProject.pdf.

10.1 Site Cohort Description Form
10.2 Data Use Agreement
10.3 Account Request Form

11.0 PROPERTY OF THE DATA BASE AND PUBLICATION POLICY

Each institution will retain full access and publishing rights to its own data; however, the collective database will be the property of the IASLC, and CRAB will be responsible for its management, storage, and analysis.

Publications related to the objectives of the Lung Cancer Staging Project of the IASLC SPFC (i.e., publications providing recommendations for changes in the TNM classification) will be planned, researched, analysed, and written by the members of the respective Subcommittees, and will follow the same authorship pattern used for the publications on the retrospective data: chairperson of the subcommittee, members of the subcommittee in alphabetical order, Chairman of the Staging and Prognostic Factors Committee, on behalf of the IASLC Staging Committee, and participating institutions.

12.0 DATA COLLECTION PROCESSES

Identification and training of data collectors will be left to the discretion of the participating institution. With the exception of the outcome data, most of the data collected for this study will occur around the time of diagnosis and treatment. The last date of follow up and vital status of each study subject will be updated at each follow-up visit with a frequency of no less than once per year.

Institutions approved by the SPFC for participation in the project will enter the data online using the secure, web-based EDC system or transfer data from an existing database. Designed and administered by Cancer Research and Biostatistics (CRAB), the system will incorporate extensive, between-field logic checks and provide a query system enabling communication between CRAB and the institutions regarding the data. The system will provide users the ability to download all data entered by that institution.

Transfer of existing, external data will be initially limited to selected partners from the retrospective project and centers that facilitate correction of geographical gaps identified in the retrospective data. Additional sites may be recruited to meet the accrual goals, provided standards regarding data quality and completeness are met.
It is the intent of the project to follow each subject until death, provided there is sufficient funding to maintain this follow-up. As of the date of activation of this protocol, the IASLC has agreed to sponsor collection of data via the EDC through the year 2024.

13.0 OVERSIGHT BY VALIDATION AND METHODOLOGY COMMITTEE

The SPFC Validation and Methodology Committee will monitor population coverage, losses to follow up, and missing data rates at each site and report their findings to the full committee.

14.0 SECONDARY USE OF THE IASLC LUNG CANCER DATA BASE

The IASLC SPFC has a duty to ensure that the data within its database are used to maximum benefit for the good of patients and the lung cancer community, within ethical constraints and the agreements entered into with individual databases. All requests for the secondary use of the database will be subjected to the following review mechanism: An initial, outline proposal should be submitted to the chair of the committee. This will be reviewed by e-mail by a sub-committee consisting of the chair person, a CRAB member of the committee, and the chair of the relevant sub-committee. If the request is considered to be a reasonable proposal, the applicant will be asked to submit a full application containing the following, additional documents:

a) A full proposal setting out the details of the study, methods, population under study, data required from the database and proposed time lines.
b) A full list of the participants to the study and proposals for involvement by members of the committee and CRAB. The study should include as primary authors at least one medical member of the committee and one CRAB member of the committee.
c) A supportive letter from CRAB confirming that the necessary data is obtainable from the data base and that the quality and volume of that data is adequate to answer the question posed.
d) Confirmation that the applicant and all other parties who may be considered to hold intellectual property rights will adhere to the highest scientific and ethical standards, including but not exclusively:
   i. Will respect the IASLC ownership of the data and will not seek to use the information provided for any other use without the agreement of the IASLC.
   ii. Will respect the anonymity of the clinical data.
   iii. Will submit any publication or presentation for scrutiny by the committee, and in addition, by those database proprietors with whom there exists prior agreements, before submission. The committee reserves the right to deny publication in extreme situations.
   iv. Will publish any submission in a format agreed with the committee, including the format of the title, and acknowledging the participation of the IASLC, the committee members, CRAB and the database proprietors. The
acknowledgment of our sponsors will be recognized in a format agreed with them from time to time.

v. Will submit publications, in the first place, to the Journal of Thoracic Oncology, the official journal of the IASLC.

The full proposal will be circulated to the full committee by e-mail and the committee’s view collected by the chairman. If consensus is not reached the proposal will be discussed at the next meeting of the committee. Revisions or additional material may be requested before a final decision is reached. The committee’s decision is final and there will be no appeal structure.
15.0 REFERENCES


